I. R. IRAN

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

International Journal of Engineering

Journal Homepage: www.ije.ir

MERC

TRANSACTIONS B: APPLICATIONS

Volume 37, Number 02, February 2024

Materials and Energy Research Center

In the Name of God

INTERNATIONAL JOURNAL OF ENGINEERING
Transactions B: Applications

DIRECTOR-IN-CHARGE
H. Omidvar

EDITOR IN CHIEF
G. D. Najafpour

ASSOCIATE EDITOR
A. Haerian

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

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

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

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

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

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

International Journal of Engineering, Volume 37, Number 02, February 2024

CONTENTS

Transactions B: Applications

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

International Journal of Engineering, Volume 37, Number 02, February 2024

IJE TRANSACTIONS B: Applications Vol. 37 No. 02, (February 2024) 213-223

International Journal of Engineering

Journal Homepage: www.ije.ir

Effect of Current Density on Magnetic and Hardness Properties of Ni-Cu Alloy Coated on Al via Electrodeposition

C. Rosyidan[a,b], B. Kurniawan*[a], B. Soegijono[c], V. G. Vidia Putra[d], D. R. Munazat[a], F. B. Susetyo[e]

[a] *Department of Physics, Universitas Indonesia, Depok, Indonesia*
[b] *Department of Petroleum, Universitas Trisakti, Jakarta, Indonesia*
[c] *Department of Geoscience, Universitas Indonesia, Depok, Indonesia*
[d] *Plasma and Nanomaterial Research Group, Politeknik STTT Bandung, Bandung, Indonesia*
[e] *Department of Mechanical Engineering, Universitas Negeri Jakarta, Jakarta, Indonesia*

PAPER INFO

Paper history:
Received 28 July 2023
Received in revised form 08 Septemebr 2023
Accepted 30 September 2023

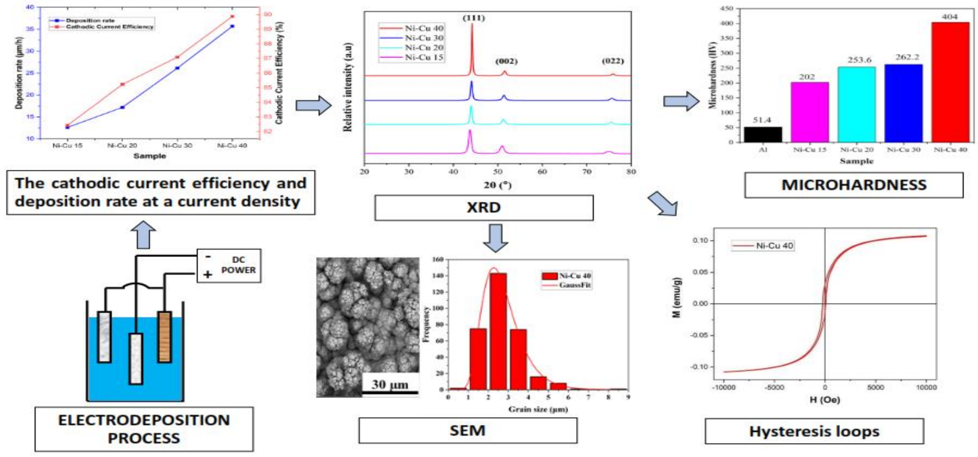
Keywords:
Cathodic Current Efficiency
Microhardness
Ni-Cu Coating
Vibrating Sample Magnetometer

ABSTRACT

Nickel (Ni)-rich single-phase nickel-copper (Ni-Cu) alloy coatings were produced on aluminum (Al) substrates by electrodeposition in stabilized citrate baths. Electrodeposition experiments were performed at four different current densities. Increasing the current density resulted in the metal deposition rate increasing faster than the hydrogen evolution rate; thus, the cathodic current efficiency increased. The crystal systems of the Ni-Cu alloys were face center cubic (fcc), with the (111) plane as the preferred crystal plane. Scanning electron microscopy with energy dispersive X-ray spectroscopy (SEM-EDS) measurements showed that the Ni content in the coating increased with increasing current density. The Ni-Cu 40 sample had the most Ni content and showed a homogeneous and compact morphology. It was found that the higher the concentration of Ni in the solution, the smaller the grain size. Measurements recorded with a vibrating sample magnetometer (VSM) showed that the Ni-Cu 40 sample provided magnetic saturation, with the highest value being 0.108 emu/g. The microhardness method produced 404 HV on the Ni-Cu 40 sample. In conclusion, higher current densities were associated with a higher Ni composition and increased thickness, which were responsible for the increases in the magnetic properties and hardness.

doi: *10.5829/ije.2024.37.02b.01*

Graphical Abstract



---

*Corresponding author email: budhy.kurniawan@sci.ui.ac.id. (B. Kurniawan)

Please cite this article as: Rosyidan C, Kurniawan B, Soegijono B, Vidia Putra VG, Munazat DR, Susetyo FB. Effect of Current Density on Magnetic and Hardness Properties of Ni-Cu Alloy Coated on Al via Electrodeposition. International Journal of Engineering, Transactions B: Applications. 2024;37(02):213-23.

NOMENCLATURE

C_e	Cathodic current efficiency	W_i	Initial weight of the substrate
σ	Lattice strain	W_f	Weight
W_s	Final weight of the substrate	I	Total current
W_m	Ratio of the final weight of the substrate	t	Deposition time
μ	Texture coefficient of the unique plane	F	Faraday's constant
$I(hkl)$	Measured intensity	f_{ni}	Nickel deposit weight ratio
m_{cu}	Copper's atomic weight	m_{ni}	Nickel's atomic weight

1. INTRODUCTION

Researchers widely study nickel (Ni) and copper (Cu) alloys as engineering materials due to their unique mechanical, magnetic, and anti-corrosion properties (1, 2). Ni-Cu alloys are known as monel in the industry and are typically comprised of 70 wt% Ni and 30 wt% Cu (3). These alloys have outstanding capabilities in acidic and alkaline environments (4). Ni-Cu alloys are single-phase alloys throughout their composition on the phase diagram, and these alloys formed because Ni and Cu are fully soluble in their solid and liquid states (5). Ni and Cu both have a face center cubic (fcc) crystal structure, and they have almost similar electronegativity and atomic radii (6, 7).

Given that conventional casting as a manufacturing method for monel results in substantial production costs and that Ni-Cu alloy coated on aluminum (Al) has potential as a replacement for monel as a bulk material (8), Al-based metals have received considerable attention (9, 10). They are lightweight and demonstrate high resistance to wear and corrosion and a high strength–stiffness combination (11).

Several techniques have been proposed to successfully modify the surface morphology and chemical composition, including sol-gel, chemical etching, chemical vapor deposition, thermal embossing, and electrodeposition (12, 13). The electrodeposition technique is a cost-effective, scalable, and easy-to-control process for coating Ni-Cu alloy (14). Specific methods have also been developed to determine the structure, morphology, and phase composition of the coated Ni-Cu alloys (15, 16). Goranova et al. (17) investigated how changing the concentration of Ni ions and the current density affected the structure and composition of Ni-Cu alloys formed by electrodeposition in alkaline citrate baths. Higher concentrations of Ni ions in the bath led to notably smoother deposits and enhanced current efficiency. However, producing a uniform Ni-Cu coating can be challenging due to the difference in reduction potential between Ni and Cu. The reduction potential of Ni atoms is -0.25 V vs. SHE, and that of Cu atoms is +0.34 V vs. SHE (18). As a result, controlling the concentrations of Ni and Cu is vital. Complexing agents must be added to narrow the potential difference between Ni and Cu. The most frequently used complexing agent is citrate due to its low toxicity, cost-effectiveness, and buffering characteristics (19).

The electrodeposition process affects the physical properties of the resultant Ni-Cu alloy, as does the current density. A high current density causes the crystal plane to be oriented in the (111) plane, the lattice size to be smaller, and the atomic distance to be less (20). The grain size becomes smaller when the current density is high, and the morphological shape resembles that of a cauliflower (21). As a result of a high current density, the coating will be thicker, and the composition of the Ni weight fraction will also be higher. The amount of Ni deposited on the substrate and the thickness of the coating both have an impact on the product's magnetic properties (22). In addition, a smaller grain size results in an increase in hardness (23). Karunakaran et al. (24) reported a hardness of 153 HV when the current density was 40 mA/cm^2, and Karunakaran and Pugazh (25) Vadivu reported a magnetic saturation value of 0.0004 emu/g at 40 mA/cm^2. Nevertheless, the researchers did not examine the impact of the coating electrodeposition factors, structure, and morphology on the magnetic and hardness properties.

The aims of this research were 1) to produce a Ni-rich Ni-Cu alloy coating on Al via electrodeposition and 2) to investigate the link between magnetic and hardness properties and the coating's microstructure and surface morphology. We varied the current density, and the process was conducted at room temperature. We examined the influence of various process variables on the cathodic current efficiency, structure, morphology, composition, grain size, and thickness of the produced coatings. Finally, the magnetic properties and hardness of the coatings were investigated.

2. MATERIAL AND METHODS

2. 1. Material and Electrodeposition Process

The chemical composition of the Al substrate (cathode) used was Fe = 1.63 wt%, Mg = 1.49 wt%, and Al = 96.88 wt%. The chemical composition of the Ni (anode) used was Al = 0.02 wt%, Ca = 0.04 wt%, Fe = 0.23 wt%, Y = 1.61 wt%, Zr = 0.04 wt%, Nb = 0.05 wt%, and Ni = 98.01 wt%. The chemical composition of the Cu (anode) used was P = 0.22 wt%, Cd = 0.684 wt%, Si = 0.137 wt%, and Cu = 98.959 wt%. The Al was cleaned from the oxide coating with sandpaper before deposition using DELTA D68H for 5 min. Ni-Cu electrodeposition was carried out using a SANFIX 305 E DC power supply. The samples produced using a current density of

15 mA/cm^2, 20 mA/cm^2, 30 mA/cm^2, and 40 mA/cm^2 were designated as Ni-Cu 15, Ni-Cu 20, Ni-Cu 30, and Ni-Cu 40, respectively. Table 1 summarized the bath composition and deposition parameters.

2. 2. Characterization The deposition rate was calculated using the previously reported method (26). The following formula, Equation 1, was used to calculate the efficiency of the cathodic current (27):

$$C_e = {W_m}/{W_f} \qquad (1)$$

W$_m$ and W$_f$ were calculated using Faraday's law, as shown in Equations 2 and 3.

$$W_m = W_s - W_i, \qquad (2)$$

$$W_f = \{(m_{ni}/2) * f_{ni} + (m_{cu}/2) * f_{cu}\} * I.\frac{t}{F}. \qquad (3)$$

The crystal structure of the Ni-Cu coating was determined using X-ray diffraction (XRD-PANanalytical Aeris Instrument Suit) (Cu-Kα radiation, λ = 0.15418 nm). XRD data were collected from 20° to 80° with a step size of 0.020°. The Materials Analysis Using Diffraction (MAUD) program was used to determine the crystal parameters of the sample after Rietveld refinement. The preferential crystallite orientation was determined from the texture coefficient μ, as shown in Equation 4 (28):

$$\mu = \frac{I(hkl)/I_0(hkl)}{(\frac{1}{N})\sum[I(hkl)/I_0(hkl)]}, \qquad (4)$$

Based on XRD results, the lattice strain σ was calculated using Equation (5) (29):

$$\sigma = \left(\frac{\beta}{4 \times \tan\theta}\right) \qquad (5)$$

SEM-EDS (Thermofisher Quanta 650 EDAX EDS Analyzer) with 1000× magnification was used to analyze the surface morphology of the Ni-Cu coatings. EDS was used to determine the chemical composition of the coatings, and the statistical distribution of grain sizes was calculated using ImageJ software. The cross sections of the coated samples were also examined to assess how the current density and CCE affected the coating thickness. Measurement of magnetic properties was conducted using a vibrating sample magnetometer (VSM, Oxford 1.2H). The hardness of the Ni-Cu coatings was measured using a MicroMct® 5100 Series Microindentation Hardness Tester. The ATM E384 standard was used for the tests, which were performed with a load of 100 g for 10 s at five places.

TABLE 1. Bath composition and deposition parameters

Bath composition and condition	Quantity
$NiSO_4.6H_2O$ (Merck)	0.5 M
$CuSO_4.5H_2O$ (Merck)	0.04 M
$Na_3C_6H_5O_7$ (Merck)	0.2 M
pH	4.2
Temperature	25 °C
Deposition time	1 h

3. RESULTS AND DISCUSSION

3. 1. Cathodic Current Efficiency and Deposition Rate Figure 1 depicts the relationship between the average CCE and deposition rate in the citrate electrolyte bath. The CCE was found to be high, with a value of 82–89%. The basic concept of current efficiency can be understood as the fraction of total current used for metal plating (30). Apart from metal deposition, hydrogen evolution is the only other necessary process that must occur on the substrate surface. Ni and Cu precipitation are both antagonistic to the hydrogen evolution reaction. In this study, the metal deposition rate increased faster than the hydrogen evolution rate when the current density increased from 15 to 40 mA/cm^2; hence, the CCE increased. The highest CCE (89.96%) was associated with the Ni-Cu 40 sample, and the lowest CCE (82.55%) was associated with the Ni-Cu 15 sample. Basori et al. and Syamsuir et al. (31) found that the deposition rate and current efficiency are considered linear.

3. 2. Structural Properties Figure 2(a) depicts the XRD spectra of the Ni-Cu alloy samples produced at various current densities. According to the XRD analysis results, each Ni-Cu alloy sample consisted of a single phase with an fcc structure. The evolution of the lattice parameters of the Ni-rich (111) phase over the range of current densities is shown in Figure 2(b). The

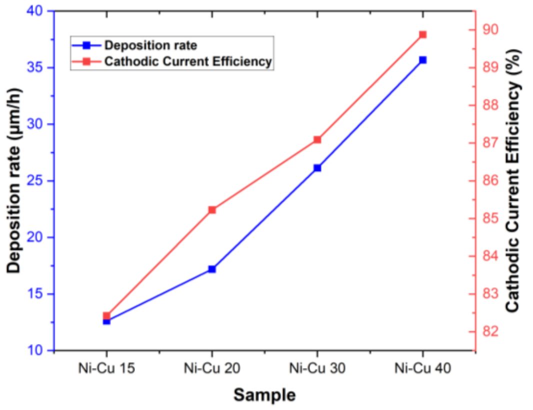
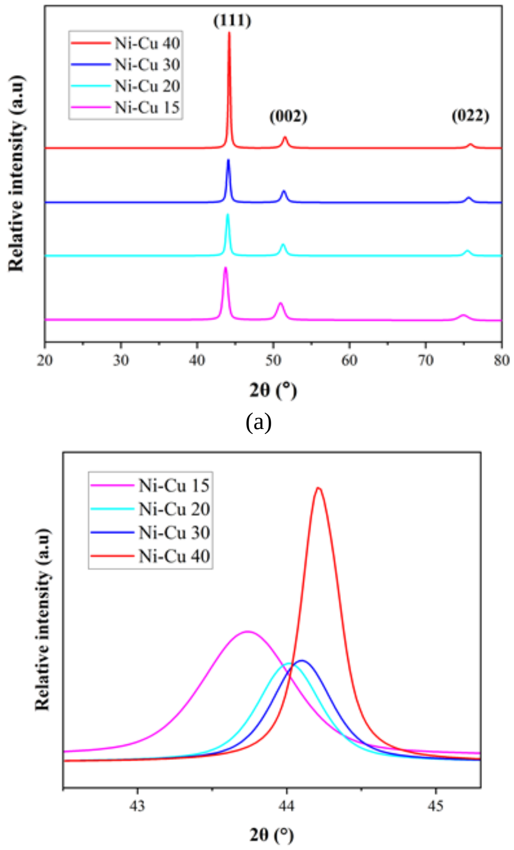


Figure 1. The cathodic current efficiency and deposition rate at a current density of 15, 20, 30, and 40 mA/cm^2



(a)

(b)

Figure 2. (a) X-ray diffraction spectra of Ni-Cu coatings electrodeposited at various current densities and (b) the extended view of the Ni-Cu (111) plane, showing peak shifts

peaks of each sample were found between the peaks of the fcc of 2θ = 43.3° for pure Cu and 2θ = 44.5 ° for pure Ni (32, 33). As the Ni content of the alloy coating increased, the diffraction angle also increased.

Crystal size calculation using MAUD resolved refinement was used to determine the size of the crystallites in the Ni-Cu alloys, and the results (Table 2) show that the crystallite size of the Ni-Cu coating ranged from approximately 24 to 50 nm. The crystallite size of a pure Ni layer is 60 nm, meaning that the Ni-Cu alloys had smaller crystallite sizes than a pure Ni layer. This result is also similar to that obtained by Li et al. (34). In contrast to the typical watt-Ni coating, we found that the presence of a sodium citrate complexing agent resulted in a finer crystallite size. This is consistent with the findings of Sarac and Baykul (35), who observed that Cu atoms affect grain refinement in Ni-Cu alloys. Cu atoms can restrain the surface diffusion of Ni atoms during the deposition process and inhibit the growth of crystallites.

The evolution of the crystallographic orientation of the Ni-Cu coatings produced with varying current densities is shown in detail in Figure 3. It can be observed that a strong (111) fiber texture appeared in all the samples, while the (002) texture gradually decreased as the current density increased.

The μ values of different crystal planes are also used to evaluate the degree of crystallographic orientation (34). Moreover, the Ni-Cu coating electrodeposited at the current density of 40 mA/cm^2 was found to have a strong (111) texture.

The texture coefficient for every preference was calculated using Equation 4 to ascertain the preferred crystal orientation direction of each Ni-Cu alloy obtained at the various current densities, and the results are shown in Table 3 (32).

It seems that the texture coefficient was also dependent on the peak current, and the preferred orientation was the (111) plane. Li et al. (34) found that the higher the current density, the more dominant the (111) plane. The findings suggest that the (111) crystallographic orientation was preferable for all the Ni-Cu coatings electrodeposited at the tested current densities.

The lattice strain of the prepared coatings was determined using Equation 5, and Figure 4 illustrates the changes in the crystal size and lattice strain of the Ni-Cu coatings based on the current density applied in the plating bath.

The crystal size increased and the lattice strain decreased as the current density increased (36). This result aligns with that obtained by Devi et al. (33), who showed that the higher the current density, the more the

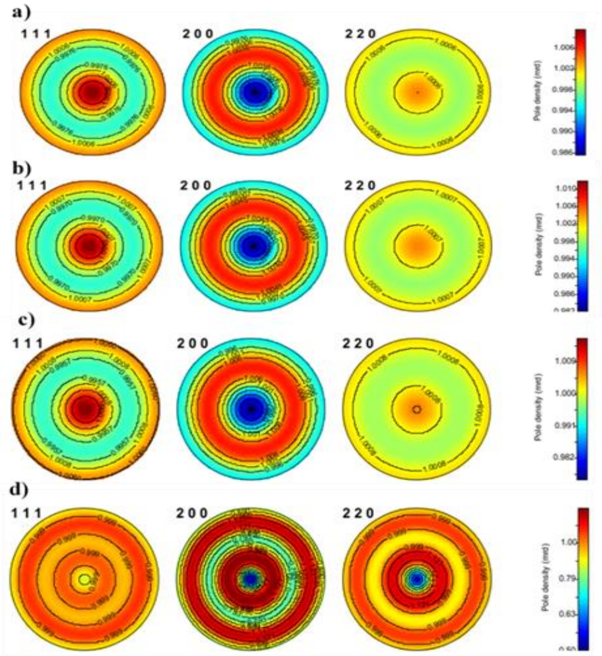


Figure 3. The simulated 2D pole figures for the (a) Ni-Cu 15, (b) Ni-Cu 20, (c) Ni-Cu 30, and (d) Ni-Cu 40 samples

TABLE 2. Parameters of the Ni-Cu alloys after Rietveld refinement using MAUD

	Sample			
Parameter	**Ni-Cu 15**	**Ni-Cu 20**	**Ni-Cu 30**	**Ni-Cu 40**
Crystal structure	Cubic fcc			
Space group	Fm-3m			
Lattice constant (Å) a = b = c	3.582	3.560	3.554	3.545
Volume (Å^3)	45.975	45.152	44.905	44.557
d-spacing (Å)	1.791	1.780	1.695	1.691
Crystallite size (nm)	24.55	26.82	32.29	50.78
Rwp (100%)	3.640	4.484	5.394	5.139
GOF	1.94	1.72	2.06	2.03
Lattice strain	0.663	0.453	0.431	0.276

TABLE 3. Texture coefficient analysis of Ni-Cu deposits

Sample	μ (hkl)		
	[111]	**[002]**	**[022]**
Ni-Cu 15	1.12	0.86	0.77
Ni-Cu 20	1.17	0.74	0.81
Ni-Cu 30	1.18	0.74	0.78
Ni-Cu 40	1.48	0.34	0.32

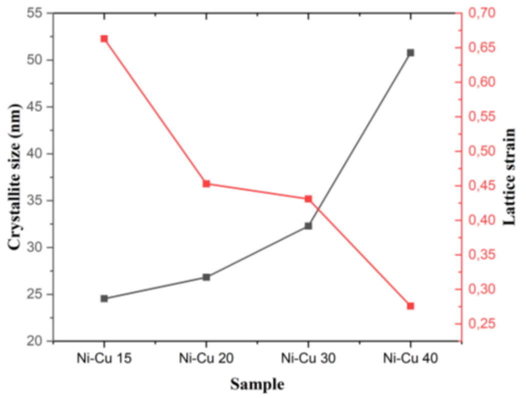


Figure 4. The lattice strain and crystallize size of the Ni-Cu alloys coated on Al at various current densities

crystal size increased. A possible reason for this is that the composition of Ni increases as the current density increases.

3. 3. Surface Morphology EDS was used to determine the elemental composition of the Ni-Cu coatings, and the results are shown in Figure 5 and Table 4. Cu and Ni were the only elements present in the deposits. The alloy composition was influenced by the current density: as the current density increased, the Cu content decreased. This phenomenon can be caused by the [Ni^{2+}]/[Cu^{2+}] ratio in the bath, which changes the composition of Ni and Cu. Goranova et al. discovered that as the Cu content of deposits decreased, so did the CCE (37). This phenomenon occurs because of the orderly deposition of Ni and Cu (38). In regular deposition, increasing the current density leads to an increased proportion of less noble metals in the deposited material (39). In our scenario, Ni is the less noble metal. As a result, when the current density was higher, the deposits were richer in Ni.

Another notable feature was that as the current density increased, the peak shifted to the right (i.e., to higher 2θ values). The change in the alloy composition may have also caused this peak shift. The Ni concentration increased as the current density increased (see Table 4). Because Ni and Cu combine to produce a single-phase alloy, the diffraction peak shifted toward that of pure Ni as the Ni percentage increased. This finding is similar to the observations of Goranova et al. (17), who found that the fcc reflection for Ni-rich Ni-Cu alloy deposits shifted to the right as the Ni concentration increased. Indeed, it is logical to expect the Ni-Cu alloy peak to shift as the Ni content increase.

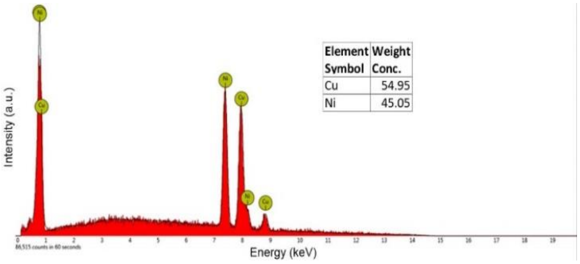


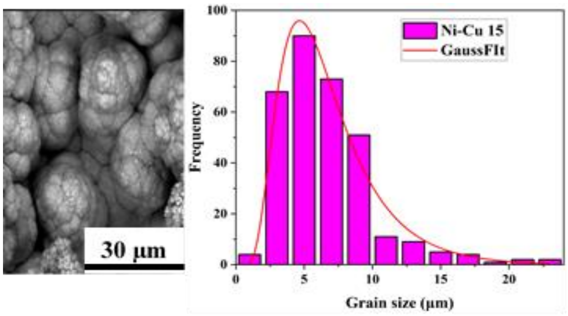
Figure 5. Eds graph of the Ni-Cu 15

TABLE 4. The chemical composition of Ni-Cu coatings on the Al substrate prepared at different current densities

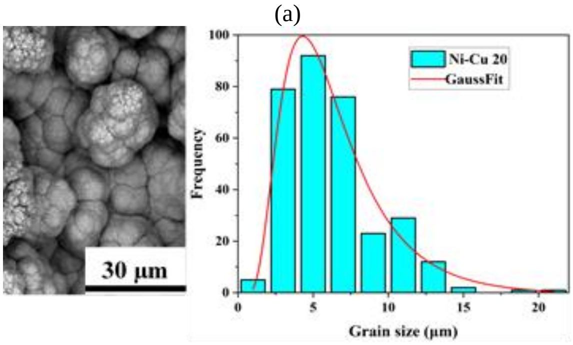
Sample	Cu, wt%	Ni, wt%
Ni-Cu 15	54.95	45.05
Ni-Cu 20	39.83	60.17
Ni-Cu 30	29.80	70.20
Ni-Cu 40	19.64	80.36

The surface morphological structure and the cross section of the coated samples were observed using SEM. Figure 6 shows SEM micrographs of the four samples' surface morphological structures. The deposits developed a fine-grained and compact spherical shape when lower deposition current densities were applied (Figure 5(a)). Deo et al. (27) and Goranova et al. (37) also observed this morphology at low current densities. The shape changed to a coarser cauliflower form when higher current densities were applied (Figure 6(d)). A diffusion-limited deposition mechanism in which a multigeneration spherical diffusion layer creates a cauliflower shape is likely to produce this type of morphology (40). As the current density increased, the cauliflower-like protrusions became more spaced and separated, creating gaps. The Ni-Cu 40 sample, produced with the highest current density, was found to have the largest gaps between the cauliflower-like bulges.

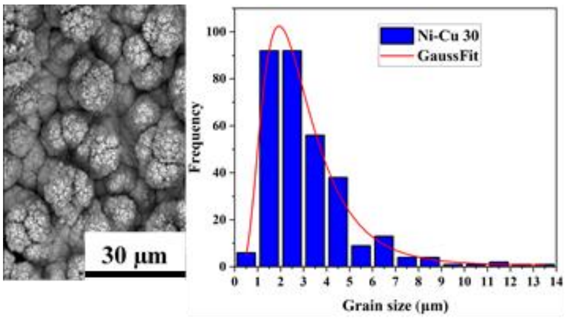
The increased nucleation rate can explain the observed decrease in grain size with increasing current density (41, 42). Ni-Cu ion flow to the cathode is faster at higher current densities. Further investigation is needed to determine the exact relationship between the current density and the grain size of the coating. In the Ni-Cu 15, Ni-Cu 20, Ni-Cu 30, and Ni-Cu 40 samples, the peak that corresponded to the (111) plane shifted toward the right (see Figure 2b). The reduction in d-spacing is ascribed to residual stress induced at a higher deposition rate (43). The statistical results of the grain size distribution presented in Figure 6 indicate that the grain size ranged from 4.63 to 1.94 µm. The decrease in grain size with the increase in current density is evident in the data shown in Table 5.



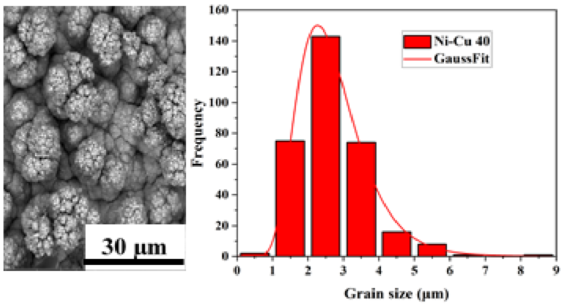
(a)



(b)



(c)



(d)

Figure 6. Surface SEM images of the deposited Ni-Cu alloy coatings and plots showing the statistical distribution of the grain size

Figure 7 depicts the relationship between the deposited alloy composition and the applied current density. EDS was used to determine the composition.

Figure 8 (a–d) displays SEM cross-section images of the produced Ni-Cu coatings. The absence of cracks between the substrate and coating demonstrates that appropriate adhesion occurred between the two entities (44). The thickness of the electrodeposited Ni-Cu coating was also measured for each sample (27), and the

TABLE 5. Average grain size found in each sample

Sample	Average grain size (µm)
Ni-Cu 15	4.63 ± 0.269
Ni-Cu 20	4.38 ± 0.365
Ni-Cu 30	2.28 ± 0.068
Ni-Cu 40	1.94 ± 0.032

following results were recorded: Ni-Cu 15 = 32 µm, Ni-Cu 20 = 42 µm, Ni-Cu 30 = 49 µm, and Ni-Cu 40 = 50 µm. The effect of the current density on the thickness of the Ni-Cu coating is depicted in Figure 8; the thickness increased as the current density increased (19). Hence, a higher current density results in more mass and a thicker coating. The findings presented in Figure 1 show that as the current density increased, so too did the CCE.

According to Faraday's law, when the deposition time remains constant for all samples, coatings formed at lower current densities will be thinner than those produced at higher current densities. A thinner covering may lead to severe interference from the Al substrate. The cathodic current density also affects the coating, as Deo et al. (27) discovered that increasing the current density improves the thickness of the film due to an increase in the CCE.

3. 4. Magnetic Properties Figure 9 depicts the magnetic characteristics and fluctuations in the current density magnetization measured with a VSM at room temperature (45). The results of the VSM analysis demonstrate that the coatings in the Ni-Cu 15, Ni-Cu 20, Ni-Cu 30, and Ni-Cu 40 samples displayed ferromagnetic activity. The low ferromagnetic activity of the $Cu_{54.95}Ni_{45.05}$ alloy film of Ni-Cu 15 could be attributed to Ni diffusion in the Cu matrix, as Cu is a diamagnetic metal and Ni is a ferromagnetic metal (46). The ferromagnetic properties of the Ni-Cu films of Ni-Cu 20, Ni-Cu 30, and Ni-Cu 40 increased with the Ni content of the alloy coatings.

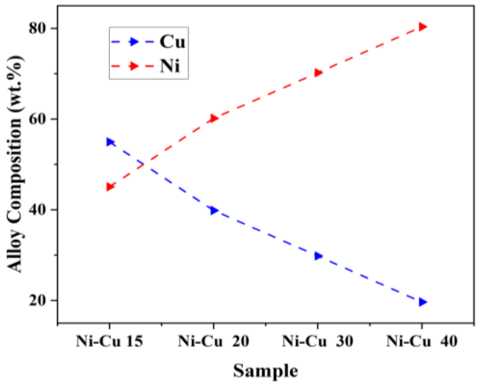


Figure 7. The dependence of the deposited alloy's composition (shown as wt% of the single electrolytes) on current density

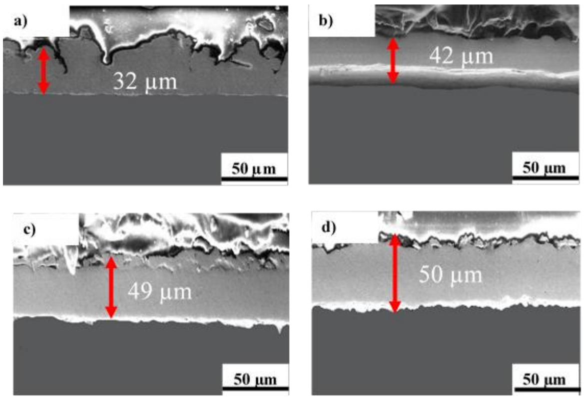
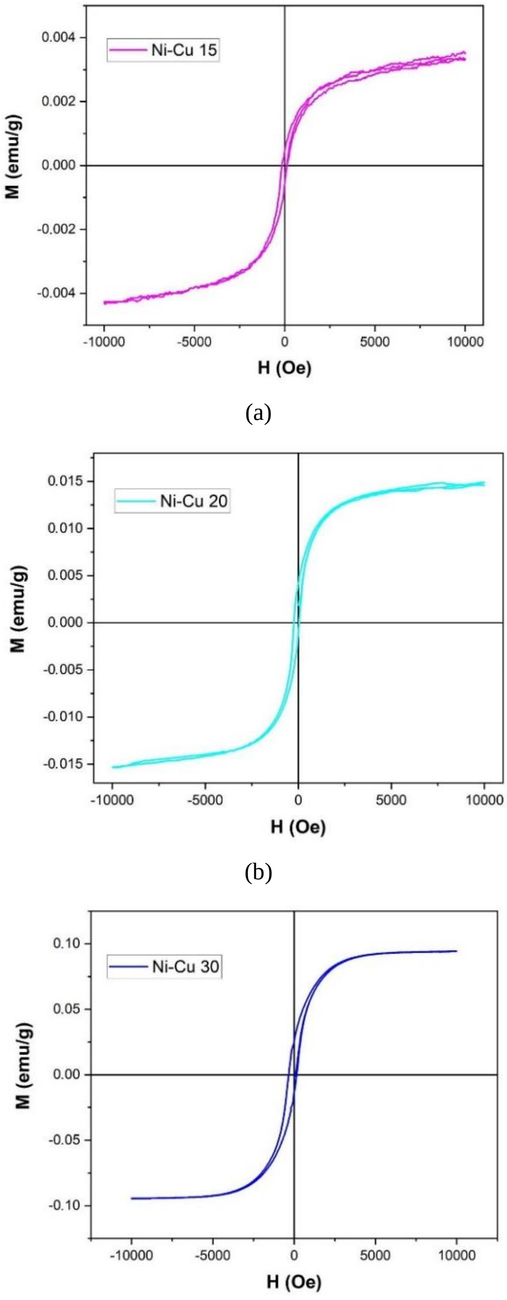


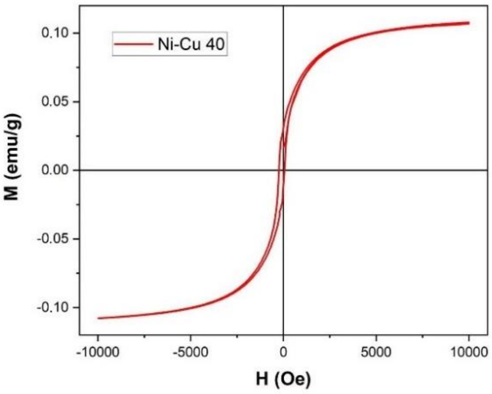
Figure 8. SEM cross-section images of the Ni-Cu coating on Al in the (a) Ni-Cu 15, (b) Ni-Cu 20, (c) Ni-Cu 30, and (d) Ni-Cu 40 samples



(a)

(b)

(c)



(d)

Figure 9. Hysteresis loops of multilayers generated at different current densities

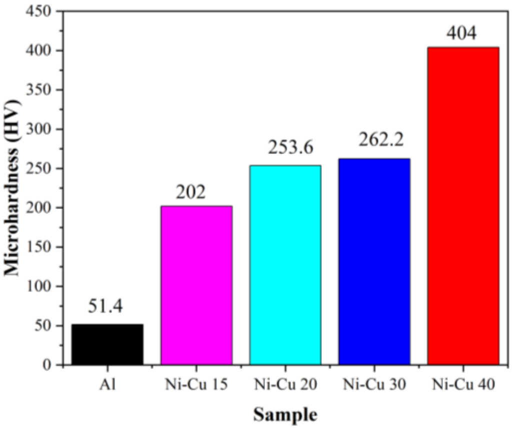


Figure 10. The microhardness of the electrodeposited Ni-Cu samples produced with different current densities

As the Ni content of the Ni-Cu alloy coatings increased, so did the saturation magnetization (see Table 6). Wang et al. (43) reported that saturation magnetization depends on the Ni content of Ni-Cu alloy coatings. In addition, Awasthi (22) reported that magnetization is enhanced by increasing the coating thickness. A possible reason for this enhanced magnetization is the magnetic disorder caused by the coating. It has been shown that the trend in saturation magnetization enhancement is associated with the coating level (47). Demidenko et al. (48) found that monel has paramagnetic properties at room temperature, while the Ni-Cu alloys in this study had ferromagnetic properties. Here, we have improved upon the results of previous studies in which phosphorus (P) and tungsten (W) were added (25). In this study, the Ni-Cu 40 sample exhibited the strongest magnetic properties.

3. 5. Hardness Figure 10 depicts the dependence of the microhardness of the Ni-Cu coatings on the current density in the plating bath. From the data presented in Figure 10, it is clear that the coating of the Ni-Cu 40 sample had the highest microhardness value (404 HV). In general, the microhardness increased with the current density and was attributed to the grain size and thickness of the coatings (42). Pingale et al. (19) found that hardness increased with the thickness of the coating. The coating thickness could influence the hardness of Ni-Cu films (26).

TABLE 6. The results of the magnetic analysis of the Ni-Cu/Al samples

Sample	Hc (Oe)	Mr (emu/g)	Ms (emu/g)
Ni-Cu 15	143.829	0.00047	0.003
Ni-Cu 20	140.081	0.004	0.015
Ni-Cu 30	256.215	0.025	0.094
Ni-Cu 40	144.023	0.032	0.108

The results indicate that alloys with a greater Ni content are mechanically harder. The overall dependence of hardness and microhardness on the percentage of Ni is shown in Figure 11, and the data indicate that microhardness increases as the percentage of Ni increase. This result is similar to Marenych's (49) finding that the hardness value is highest with the highest Ni composition.

Moreover, the hardness reported in previous studies that resulted from electrodeposition of Ni-Cu on Al in the presence of P was lower than that recorded in the present study (24). This is due to the smaller grain size that resulted from applying a different current density. In addition, the increase in microhardness reported here is related to the role that Ni atoms play in grain refinement (31). Ramkumar et al. (50) reported a monel hardness value of 165 HV, which is lower than the peak hardness value recorded in the current study.

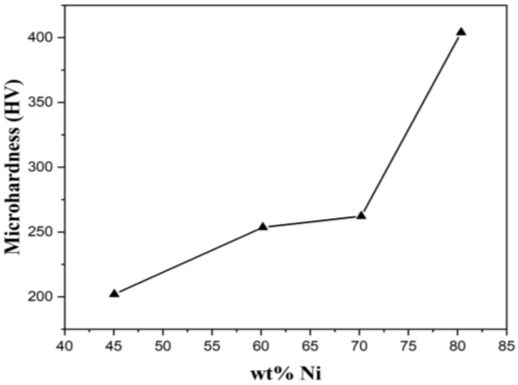


Figure 11. The relationship found between wt% Ni and microhardness, based on the data obtained from the four experimental samples

4. CONCLUSION

In this study, Ni-Cu alloys were electrodeposited onto Al substrates using citrate baths. The effects of adding a citrate solution on the properties of the deposited coatings were studied, and the results illustrate that single-phase Ni-Cu alloy layers were produced on the Al surface at all current densities. The CCE increased as the current density increased. The coatings formed at lower current densities showed a more compact and spherical morphology, while those formed at higher current densities showed a less uniform structure with a cauliflower-like morphology. Both the surface morphology and composition of the coating showed a strong dependence on the current density. The Ni-Cu alloy coating deposited at a low current density had a layer thickness of 32 μm, while the coating deposited at a high current density had a thickness of 50 μm. The saturation magnetization of the coating increased with the Ni content in the Ni-Cu alloy and with the coating thickness. The hardness increased with the coating thickness, grain size, and Ni composition in the solution. The hardness of the produced Ni-Cu alloy coatings was found to be greater than that of monel.

5. ACKNOWLEDGMENTS

The authors extend their gratitude to the Ministry of Research, Technology and Higher Education of the Republic of Indonesia for the financial support of Hibah Penelitian Disertasi Doktor No: NKB-971/UN2.RST/HKP.05.00/2022.

6. REFERENCES

1. Chen Z, Wang C, Tang C, Lek YZ, Kandukuri SY, Du H, et al. Microstructure and mechanical properties of a Monel K-500 alloy fabricated by directed energy deposition. Materials Science and Engineering: A. 2022;857:144113. https://doi.org/10.1016/j.msea.2022.144113
2. Kukliński M, Bartkowska, A., and Przestacki, D. . Microstructure and selected properties of Monel 400 alloy after laser heat treatment and laser boriding using diode laser. International Journal of Advanced Manufacturing Technology. 2018;98(9-12): 3005–17. https://doi.org/10.1007/s00170-018-2343-9
3. Mohagheghpour E, Larijani M, Rajabi M, Gholamipour R. Effect of Silver Clusters Deposition on Wettability and Optical Properties of Diamond-like Carbon Films. International Journal of Engineering, Transactions C: Aspects, . 2021;34(3):706-13. https://doi.org/10.5829/ije.2021.34.03c.15
4. Nady H, Negem M. Ni–Cu nano-crystalline alloys for efficient electrochemical hydrogen production in acid water. RSC advances. 2016;6(56):51111-9. https://doi.org/10.1039/c6ra08348j
5. Negem M, Nady H. Electroplated Ni-Cu nanocrystalline alloys and their electrocatalytic activity for hydrogen generation using alkaline solutions. international journal of hydrogen energy. 2017;42(47):28386-96. https://doi.org/10.1016/j.ijhydene.2017.09.147
6. Guisbiers G, Khanal S, Ruiz-Zepeda F, De La Puente JR, José-Yacaman M. Cu–Ni nano-alloy: mixed, core–shell or Janus nano-particle? Nanoscale. 2014;6(24):14630-5. https://doi.org/10.1039/c4nr05739b
7. Mae Y. What the Darken–Gurry plot means about the solubility of elements in metals. Metallurgical and Materials Transactions A. 2016;47(12):6498-506. https://doi.org/10.1007/s11661-016-3730-1
8. Alizadeh M, Safaei H. Characterization of Ni-Cu matrix, Al2O3 reinforced nano-composite coatings prepared by electrodeposition. Applied Surface Science. 2018;456:195-203. https://doi.org/10.1016/j.apsusc.2018.06.095
9. Toghraei M, Siadati H. Electrodeposited co-pi catalyst on α-Fe2O3 photoanode for water-splitting applications. International Journal of Engineering. 2018;31(12):2085-91. 10.5829/ije.2018.31.12c.13
10. KK P. Experimental investigation by cryogenic treatment of aluminium 6063 and 8011 and nicow coating to improve hardness and wear. International Journal of Engineering. 2016;29(6):827-33. https://doi.org/10.5829/idosi.ije.2016.29.06c.12
11. Kumar D, Angra S, Singh S. Mechanical properties and wear behaviour of stir cast aluminum metal matrix composite: a review. International Journal of Engineering. 2022;35(4):794-801. https://doi.org/10.5829/IJE.2022.35.04A.19
12. Moosaei H, Zareei A, Salemi N. Elevated Temperature Performance of Concrete Reinforced with Steel, Glass, and Polypropylene Fibers and Fire-proofed with Coating. International Journal of Engineering. 2022;35(5):917-30. https://doi.org/10.5829/ije.2022.35.05b.08
13. Poursaeidi E, Salarvand A. Comparison of properties of ti/tin/ticn/tialn film deposited by cathodic arc physical vapor and plasma-assisted chemical vapor deposition on custom 450 steel substrates. International Journal of Engineering. 2016;29(10):1459-68. https://doi.org/10.5829/idosi.ije.2016.29.10a.17
14. Allahyarzadeh M, Aliofkhazraei M, Rouhaghdam AS, Torabinejad V. Gradient electrodeposition of Ni-Cu-W (alumina) nanocomposite coating. Materials & Design. 2016;107:74-81. https://doi.org/10.1016/j.matdes.2016.06.019
15. Geramipour F, Khoei SM, Gugtapeh HS. Effect of shaped waveform on structure and electrochemical corrosion behavior of pulse electrodeposited NiCu alloy coatings. Surface and Coatings Technology. 2021;424:127643. https://doi.org/10.1016/j.surfcoat.2021.127643
16. Thurber CR, Ahmad YH, Sanders SF, Al-Shenawa A, D'Souza N, Mohamed AM, et al. Electrodeposition of 70-30 Cu–Ni nanocomposite coatings for enhanced mechanical and corrosion properties. Current Applied Physics. 2016;16(3):387-96. https://doi.org/10.1016/j.cap.2015.12.022
17. Goranova D, Rashkov R, Avdeev G, Tonchev V. Electrodeposition of Ni–Cu alloys at high current densities: details of the elements distribution. Journal of Materials Science. 2016;51:8663-73. https://doi.org/10.1007/s10853-016-0126-y
18. Lee W-H, Chung K. Investigation of a copper–nickel alloy resistor using co-electrodeposition. Journal of Applied Electrochemistry. 2020;50:535-47. https://doi.org/10.1007/s10800-020-01398-0
19. Pingale AD, Belgamwar SU, Rathore JS. Effect of graphene nanoplatelets addition on the mechanical, tribological and corrosion properties of Cu–Ni/Gr nanocomposite coatings by electro-co-deposition method. Transactions of the Indian Institute of Metals. 2020;73:99-107. https://doi.org/10.1007/s12666-019-01807-9

20. Hughes RA, Menumerov E, Neretina S. When lithography meets self-assembly: a review of recent advances in the directed assembly of complex metal nanostructures on planar and textured surfaces. Nanotechnology. 2017;28(28):282002. https://doi.org/10.1088/1361-6528/aa77ce
21. Kamel M, Anwer Z, Abdel-Salam I, Ibrahim I. Electrodeposition of nanocrystalline Ni–Cu alloy from environmentally friendly lactate bath. Surface and Interface Analysis. 2014;46(7):442-8. https://doi.org/10.1002/sia.5525
22. Awasthi S, Pandey SK, Balani K. Tuning the magnetism and tribological behaviour of electrodeposited Ni/Cu bi-layer by selective reinforcement of carbon nanotubes. Journal of Alloys and Compounds. 2020;818:153287. https://doi.org/10.1016/j.jallcom.2019.153287
23. Heidarzadeh A, Saeid T. Correlation between process parameters, grain size and hardness of friction-stir-welded Cu–Zn alloys. Rare Metals. 2018;37:388-98. https://doi.org/10.1007/s12598-016-0704-9
24. Karunakaran M, Pugazhvadivu M, Gunasegaran V, Gowtham G. Electrodeposition of Cu-Ni-PW Composite on Al-6063 Substrate. 2018. https://doi.org/10.26438/ijsrpas/v6i3.5964
25. Karunakaran M, Vadivu MP. Magnetic and micro-mechanical behavior of Cu-Ni-PW-TiO2 hybrid composite electroplating on Al alloy substrate. Journal of Magnetism and Magnetic Materials. 2019;475:359-67. https://doi.org/10.1016/j.jmmm.2018.11.077
26. Soegijono B, Susetyo F, editors. Magnetic field exposure on electroplating process of ferromagnetic nickel ion on copper substrate. Journal of Physics: Conference Series; 2022: IOP Publishing.
27. Deo Y, Guha S, Sarkar K, Mohanta P, Pradhan D, Mondal A. Electrodeposited Ni-Cu alloy coatings on mild steel for enhanced corrosion properties. Applied Surface Science. 2020;515:146078. https://doi.org/10.1016/j.apsusc.2020.146078
28. Seakr R. Microstructure and crystallographic characteristics of nanocrystalline copper prepared from acetate solutions by electrodeposition technique. Transactions of Nonferrous Metals Society of China. 2017;27(6):1423-30. https://doi.org/10.1016/S1003-6326(17)60164-X
29. Dolabella S, Borzì A, Dommann A, Neels A. Lattice strain and defects analysis in nanostructured semiconductor materials and devices by high-resolution X-ray diffraction: theoretical and practical aspects. Small Methods. 2022;6(2):2100932. https://doi.org/10.1002/smtd.202100932
30. Budi S, Tawwabin RA, Cahyana U, Paristiowati M. Saccharin-assisted galvanostatic electrodeposition of nanocrystalline FeCo films on a flexible substrate. International Journal of Electrochemical Science. 2020;15(7):6682-94. https://doi.org/10.20964/2020.07.74
31. Syamsuir S, Soegijono B, Yudanto SD, Basori B, Ajiriyanto MK, Nanto D, et al. Electrolyte Temperature Dependency of Electrodeposited Nickel in Sulfate Solution on the Hardness and Corrosion Behaviors. International Journal of Engineering, Transactions C: Aspects, . 2023;36(6):1193-200. 10.5829/ije.2023.36.06c.18
32. Soegijono B, Susetyo FB, Fajrah MC. Electrodeposition of Paramagnetic Copper Film under Magnetic Field on Paramagnetic Aluminum Alloy Substrates. e-Journal of Surface Science and Nanotechnology. 2020;18:281-8. https://doi.org/10.1380/EJSSNT.2020.281
33. Devi C, Ashokkumar R. INFLUENCE OF DIFFERENT CURRENT DENSITY ON CHARACTERISTICS OF NiFeP NANO ALLOY THIN FILMS. Rasayan Journal of Chemistry. 2018;11(3). https://doi.org/10.31788/RJC.2018.1133088
34. Li B, Mei T, Li D, Du S. Ultrasonic-assisted electrodeposition of Ni-Cu/TiN composite coating from sulphate-citrate bath: Structural and electrochemical properties. Ultrasonics sonochemistry. 2019;58:104680. https://doi.org/10.1016/j.ultsonch.2019.104680
35. Sarac U, Baykul MC. Morphological and microstructural properties of two-phase Ni–Cu films electrodeposited at different electrolyte temperatures. Journal of alloys and compounds. 2013;552:195-201. https://doi.org/10.1016/j.jallcom.2012.10.071
36. Arasteh J. Microhardness Optimization of Al–TiC Nanocomposite Produced by Mechanical Milling and Heat Treatment. Advanced Ceramics Progress. 2021;7(1):35-45. https://doi.org/10.30501/ACP.2021.265197.1052
37. Goranova D, Avdeev G, Rashkov R. Electrodeposition and characterization of Ni–Cu alloys. Surface and Coatings Technology. 2014;240:204-10. https://doi.org/10.1016/j.surfcoat.2013.12.014
38. Hu G, Huang R, Wang H, Zhao Q, Zhang X. Facile galvanic replacement deposition of nickel on copper substrate in deep eutectic solvent and its activation ability for electroless Ni–P plating. Journal of Solid State Electrochemistry. 2022;26(5):1313-22. https://doi.org/10.1007/s10008-022-05172-4
39. Niu J, Song M, Zhang Y, Zhang Z. Dealloying induced nanoporosity evolution of less noble metals in Mg ion batteries. Journal of Magnesium and Alloys. 2021;9(6):2122-32. https://doi.org/10.1016/j.jma.2021.04.003
40. Wang S, Guo X, Yang H, Dai J, Zhu R, Gong J, et al. Electrodeposition mechanism and characterization of Ni–Cu alloy coatings from a eutectic-based ionic liquid. Applied Surface Science. 2014;288:530-6. https://doi.org/10.1016/j.apsusc.2013.10.065
41. Nwaeju CC, Eboh, A. O., and Edoziuno, F. O. . Grain size evolution mechanical and corrosion behaviour of precipitate strengthened Cu-Ni alloy. Acta Metallurgica Slovaca. 2022;28(4):188–96. https://doi.org/10.36547/ams.28.4.1609
42. Ameri Ekhtiarabadi T, Zandrahimi M, Ebrahimifar H. The Impact of Current Density of Electroplating on Microstructure and Mechanical Properties of Ni-ZrO2-TiO2 Composite Coating. Advanced Ceramics Progress. 2020;6(1):22-9. https://doi.org/10.30501/acp.2020.233518.1038
43. Wang C, Hossain Bhuiyan ME, Moreno S, Minary-Jolandan M. Direct-write printing copper–nickel (Cu/Ni) alloy with controlled composition from a single electrolyte using co-electrodeposition. ACS applied materials & interfaces. 2020;12(16):18683-91. https://doi.org/10.1021/acsami.0c01100
44. Hanachi M, Seyedraoufi Z, Abouei V. Investigation of Microstructure, Hardness, and Corrosion Resistance of Ni-P-GO Electroless Nanocomposite Coating on AZ31D Alloy Surface. Advanced Ceramics Progress. 2020;6(3):55-62. https://doi.org/10.30501/acp.2020.233518.1038
45. Dhara B, Jha PK, Gupta K, Bind VK, Ballav N. Diamagnetic Molecules Exhibiting Room-Temperature Ferromagnetism in Supramolecular Aggregates. The Journal of Physical Chemistry C. 2017;121(22):12159-67. https://doi.org/10.1021/acs.jpcc.7b02145
46. Qasim I, Waqee-ur-Rehman M, Mumtaz M, Hussain G, Nadeem K, Shehzad K. Ferromagnetic (Ni) nanoparticles–CuTl-1223 superconductor composites. Journal of Magnetism and Magnetic Materials. 2016;403:60-7. https://doi.org/10.1016/j.jmmm.2015.11.066
47. Padmapriya G, Manikandan A, Krishnasamy V, Jaganathan SK, Antony SA. Enhanced catalytic activity and magnetic properties of spinel Mn x Zn 1− x Fe 2 O 4 (0.0≤ x≤ 1.0) nano-photocatalysts by microwave irradiation route. Journal of Superconductivity and Novel Magnetism. 2016;29:2141-9. https://doi.org/10.1007/s10948-016-3527-x

48. Demidenko O, Zhyvulka A, Yanushkevich K, Galias A, Constantin V, Neacsu E, et al. Magnetic properties of stainless steels under corrosive action of based on choline chloride ionic liquids. Journal of Magnetism and Magnetic Materials. 2019;477:74-6. https://doi.org/10.1016/j.jmmm.2019.01.034

49. Marenych O, Ding D, Pan Z, Kostryzhev A, Li H, van Duin S. Effect of chemical composition on microstructure, strength and wear resistance of wire deposited Ni-Cu alloys. Additive Manufacturing. 2018;24:30-6. https://doi.org/10.1016/j.addma.2018.08.003

50. Ramkumar KD, Joshi V, Pandit S, Agrawal M, Kumar OS, Periwal S, et al. Investigations on the microstructure and mechanical properties of multi-pass pulsed current gas tungsten arc weldments of Monel 400 and Hastelloy C276. Materials & Design. 2014;64:775-82. https://doi.org/10.1016/j.matdes.2014.08.055

Persian Abstract

چکیده

پوشش‌های آلیاژی تک فاز نیکل مسغنی از نیکل (Ni) بر روی بسترهای آلومینیومی (Al) با رسوب الکتریکی در حمام سیترات تثبیت شده تولید شدند. آزمایش‌های رسوب الکتریکی در چهار چگالی جریان مختلف انجام شد. افزایش چگالی جریان منجر به افزایش سرعت رسوب فلز سریعتر از نرخ تکامل هیدروژن شد. بنابراین، راندمان جریان کاتدی افزایش یافت. سیستم‌های کریستالی آلیاژهای Ni-Cu مکعبی در مرکز سطح (FCC) بودند، با صفحه (۱۱۱) به عنوان صفحه کریستالی ترجیحی. میکروسکوپ الکترونی روبشی با اندازه‌گیری‌های طیف‌سنجی پرتو ایکس پراکنده انرژی (SEM-EDS) نشان داد که محتوای نیکل در پوشش با افزایش چگالی جریان افزایش می‌یابد. نمونه Ni-Cu 40 بیشترین مقدار نیکل را داشت و مورفولوژی همگن و فشرده را نشان داد. مشخص شد که هر چه غلظت نیکل در محلول بیشتر باشد، اندازه دانه کوچکتر است. اندازه‌گیری‌های ثبت‌شده با یک مغناطیس‌سنج نمونه ارتعاشی (VSM) نشان داد که نمونه Ni-Cu 40 اشباع مغناطیسی را ارائه می‌کند، با بالاترین مقدار ۰.۱۰۸ emu/g. روش میکروسختی ٤٠٤ HV بر روی نمونه Ni-Cu 40 تولید کرد. در نتیجه، چگالی جریان بالاتر با ترکیب نیکل بالاتر و افزایش ضخامت همراه بود که مسئول افزایش خواص مغناطیسی و سختی بود.

IJE TRANSACTIONS B: Applications Vol. 37 No. 02, (February 2024) 224-236

International Journal of Engineering

Journal Homepage: www.ije.ir

Construction of Underground and Multi-story Car Parks in High-density Urban Areas

O. V. Trushko*, V. L. Trushko, P. A. Demenkov

Faculty of construction, Saint Petersburg Mining University, Saint-Petersburg, Russia

PAPER INFO

Paper history:
Received 10 July 2023
Received in revised form 23 October 2023
Accepted 05 November 2023

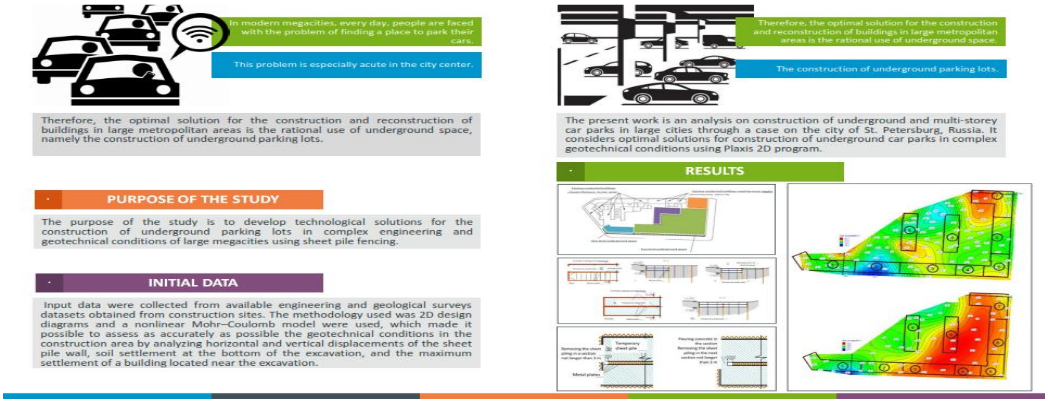
Keywords:
Construction
Metropolitan Areas
Underground Car Parks
Multi-story Car Parks
Special Methods
Sheet Piling
Complex Geotechnical Conditions

ABSTRACT

In modern megacities, every day, people are faced with the problem of finding a place to park the cars, this problem is especially acute in the city center. The number of existing parking spaces is sorely lacking in view of the rapid growth in the number of cars. Therefore, the optimal solution for the construction and reconstruction of buildings in large metropolitan areas is the rational use of underground space, namely the construction of underground parking lots. The present work is an analysis on construction of underground and multi-story car parks in large cities through a case on the city of St. Petersburg, Russia. It considers optimal solutions for construction of underground car parks in complex geotechnical conditions using Plaxis 2D program. The purpose of the study is to develop technological solutions for the construction of underground parking lots in complex engineering and geological conditions of large megacities with the use of sheet pile fencing. Input data were collected from available engineering and geological surveys datasets obtained from construction sites. The methodology used was 2D design diagrams and a nonlinear Mohr–Coulomb model was used, which made it possible to assess as accurately as possible the geotechnical conditions in the construction area by analyzing horizontal and vertical displacements of the sheet pile wall, soil settlement at the bottom of the excavation, and the maximum settlement of a building located near the excavation. As a result, professional recommendations were developed for construction of underground and multi-storey car parks in complex geotechnical conditions: it is necessary to carry out complex geotechnical support of construction; when constructing a pit in difficult engineering and geological conditions, construct a pit with metal spacer systems (open pit); when constructing a pit near existing buildings and structures, it is very important to take into account the relative position of the base of the foundation and the pit being constructed; accurately and reliably perform calculations in the design and implementation of the construction of underground parking as part of new construction or reconstruction of previously constructed buildings in difficult engineering and geological conditions. The authors of the article have developed new technological solutions that are of great scientific and practical significance to improve the reliability and safety of preserved architectural monuments during the construction of underground parking lots, as well as the safety of neighboring buildings that may be affected by construction or reconstruction.

***doi**: 10.5829/ije.2024.37.02b.02*

Graphical Abstract



In modern megacities, every day, people are faced with the problem of finding a place to park their cars.

This problem is especially acute in the city center.

Therefore, the optimal solution for the construction and reconstruction of buildings in large metropolitan areas is the rational use of underground space, namely the construction of underground parking lots.

PURPOSE OF THE STUDY

The purpose of the study is to develop technological solutions for the construction of underground parking lots in complex engineering and geotechnical conditions of large megacities using sheet pile fencing.

INITIAL DATA

Input data were collected from available engineering and geological surveys datasets obtained from construction sites. The methodology used was 2D design diagrams and a nonlinear Mohr–Coulomb model were used, which made it possible to assess as accurately as possible the geotechnical conditions in the construction area by analyzing horizontal and vertical displacements of the sheet pile wall, soil settlement at the bottom of the excavation, and the maximum settlement of a building located near the excavation.

Therefore, the optimal solution for the construction and reconstruction of buildings in large metropolitan areas is the rational use of underground space.

The construction of underground parking lots.

The present work is an analysis on construction of underground and multi-storey car parks in large cities through a case on the city of St. Petersburg, Russia. It considers optimal solutions for construction of underground car parks in complex geotechnical conditions using Plaxis 2D program.

RESULTS

---

*Corresponding Author Email: *Trushko_OV@pers.spmi.ru* (O. V. Trushko)

Please cite this article as: Trushko OV, Trushko VL, Demenkov PA. Construction of Underground and Multi-story Car Parks in High-density Urban Areas. International Journal of Engineering, Transactions B: Applications. 2024;37(02):224-36.

1. INTRODUCTION

In countries with advanced economies, underground parking is gaining more and more popularity due to the lack of space on the surface and the growing number of cars. Some cities in Asia already include in their city policy the presence of underground garages under certain standards under all new buildings.

In Europe, one of the deepest car parks in Leiden, the Netherlands, is about 22 meters deep, which corresponds to a 7-storey building. In China, in the city of Hangzhou, an automatic 12-level par-king with a depth of about 40 meters for 200 cars was built. Obviously, with the growth in the number of cars, underground parking lots will only deepen. The University of Minnesota Civil Engineering Faculty building stands out from public buildings, which has 7 underground floors, and it was built back in 1983.

Continuous global growth of urban areas and high-rise construction with underground story requires attention to lowering development footprint in order to preserve natural landscapes that makes it important to provide sufficient spaces, such as, for vegetated open areas and car parks .

Modern cities are facing major issues with acute shortage of parking spaces, especially in city centres due constant growth in the number of cars. On the other hand, while parked vehicles impose negative impacts on lawns, grass verges and general environmental situation, double-parking constructions also occupy spaces for cars to pass by and negatively influence the road safety.

Therefore, efficient use of underground spaces, including development of underground car parks, deems to provide optimal solution for construction and renovation of buildings in large cities. However, a range of difficulties commonly arise in underground parking construction, such as, due to high density in city centers and complex geotechnical conditions (1-8).

Figure 1 shows the past twenty years history of investment and leading countries, including China, the Republic of Korea, Japan, the USA, France, Taiwan, Canada, and Germany, in development of innovative technologies for underground parking construction.

Note that Intelligent Polytron Technologies (China), Beijing Zhongyan Zhibo Technology (China) and Hangzhou Haoche Technology (China) are major companies that lead in the development of innovations for underground parking construction are shown in Figures 2 and 3.

Over a long period of operation of residential buildings, the requirements for living conditions have also changed, which led to obsolescence of residential buildings of old buildings. The elimination of factors of physical and moral depreciation of residential buildings of old buildings occurs during their overhaul and reconstruction. In the case of cultural heritage objects, this is an adaptation for modern use. A demanded and justified action to improve the investment attractiveness and quality of life of citizens living in the central districts of the city is underground parking, the construction of which should create zones of comfortable stay by eliminating car parking in courtyards and on adjacent streets (9, 10).

Today, underground parking lots are an essential element of the infrastructure of any urban facility: a shopping mall, business center or residential building. The cost of real estate directly depends on the characteristics and capacity of the parking lot.

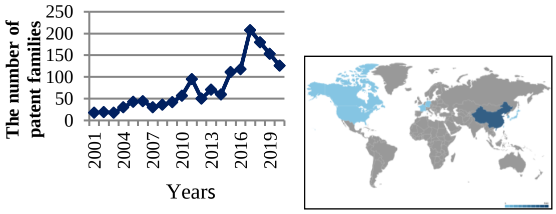


Figure 1. Investment trends over the past twenty years (Left) and Leading countries (Right) in underground parking construction technologies

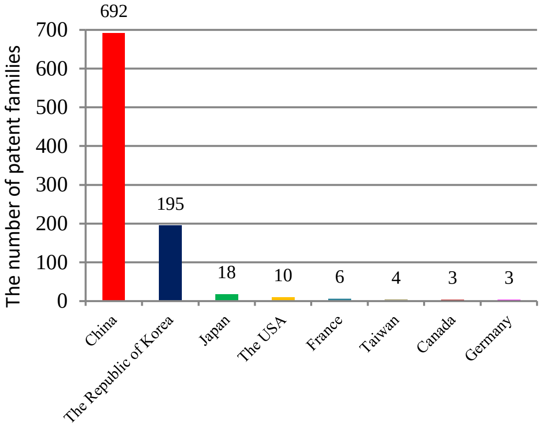


Figure 2. Patent families in the field of underground parking construction by country

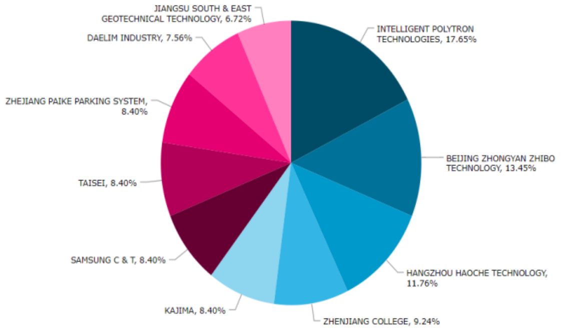


Figure 3. Shares of key players in the development of innovative technologies for underground parking construction

The advantages of underground parking lots are obvious: they save free space, can be located under roads and buildings, and sanitary and hygienic requirements for their location are much milder compared to the norms for surface parking lots and garages. Underground parking lots also save energy by reducing energy consumption due to the constant temperature of the air underground (provided the building is well insulated). The disadvantage of underground parking lots is the high cost of their construction. It is necessary to take into account internal communications, hydrogeological conditions and other factors. On the other hand, designing an underground parking lot allows to effectively use the remaining area of the structure for residential, office and commercial purposes. In the development of the central districts of the city for a number of reasons should be preferred to underground parking above ground traditional format. Designing an underground parking lot requires many details to be taken into account. First of all, these include safety, ease of use, manufacturability, waterproofing, availability of necessary engineering systems, sufficient width of entrances and exits, ceiling height and much more. Often it is not easy to comply with all the requirements, and therefore the construction of underground parking should involve professionals (11, 12)

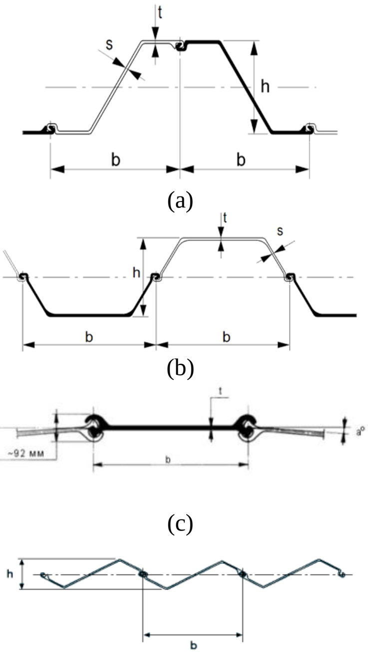
The most important characteristic of the site for building an underground parking lot is the hydrogeological conditions, in particular, water currents and soil composition. The problems of stabilization of weak soils in Russia were studied by many research scientists (13-21). Numerous works by Dashko et al. (22) are devoted to the issues of justification of building construction in complex engineering-geological and hydrogeological conditions of St. Petersburg. These features can limit the depth of the parking lot and significantly complicate construction works (13).

It is also necessary to take into account the possible impact of the new construction on the foundations of nearby buildings, as well as various urban planning restrictions.

2. DATA COLLECTION AND METHODOLOGY

Depending on the geotechnical conditions and density of urban development, countries use different innovative approaches and technologies in underground parking construction. For instance, in dense urban areas, sheet piling has become a popular choice in construction of underground car parks since it does not have a negative effect on the surrounding buildings and structures .

Sheet piles have a relatively high moment of resistance in relation to their cross-sectional area and a small thickness (Figure 4).



(a)

(b)

(c)

(d)

Figure 4. Types of sheet piling (a – Z-type sheet pile; b – U-type sheet pile; c – flat sheet pile; d- jagged wall of Z-type sheet piles)

In constructing sheet pile walls, metal profiles are most often used, which vary in their cross-sections and interlock types. Hook-type interlocks provide for ease of installation and produce a strong connection between two elements, ensuring the strength of the structure (14-16).

Sheet piles are commonly used as permanent structural elements in substructures and other applications. For example, in order to build a block of flats with four floors and one underground level in a densely populated area in France, stringent requirements were imposed on the construction site, which were partly due to its being located close to a tramway track and other buildings, including a 19th-century monastery (23-33). The allowable deflection of the sheet pile wall was limited to 20 mm and a fifty-year service life had to be guaranteed for the steel elements. To protect the nearby high-voltage power lines, a nine-metre-high barrier was built (Figure 5).

As retaining walls made of steel sheet piles take up only a minimum of space, it made them the ideal retaining-wall solution for a new underground car park. The original design included sheet piles as temporary retaining elements and a permanent concrete wall. Later, changes were made to the project to use sheet piles as permanent elements. In addition to improving the watertightness of the excavation, the sheet pile wall also resists lateral earth pressure, including traffic and tram loads.

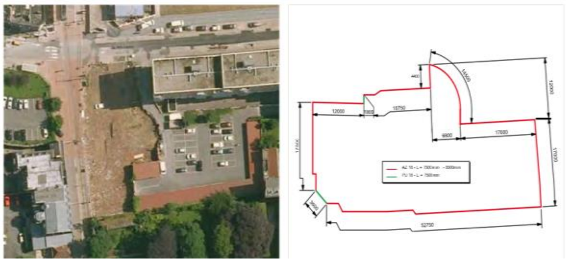


Figure 5. Plan of the construction site of the block of flats with an AZ 18 sheet pile wall that surrounded the excavation during the construction phase and was subsequently left in place to serve as the outer wall of the underground car park

The vertical loads were taken by a deep foundation of bored concrete piles. Several boreholes were drilled to a depth of up to 16 m. Soil analysis revealed the following soil layers (top to bottom):

•one metre of made ground;

•up to seven metres of silty to clayey river sand (angle of internal friction: 20-28°, cohesion up to 20 kN/m^2);

•groundwater 1.5 m below ground level (El. +40.30 m);

• hard clayey substratum down to the well bottom.

Before driving the sheet piles, the clayey soils were loosened with a 460-mm-diameter drill. The project mainly used AZ 18 profiles, as well as several PU 18 sheet piles, with both types being made of S 240 GP steel. The AZ and PU sheet piles with a length of 7.5 to 8.5 m were driven into the hard clayey substratum using a rig-mounted ABI vibratory hammer (MRZV 925).

To prevent damage to surrounding buildings, the city had introduced strict vibration limits. Therefore, special equipment was installed at the construction site to measure the vibration level.

After the sheet piles were driven into the clayey substratum, an RC capping beam was poured onto the piles.

To carry the horizontal reaction of the retaining walls during construction, large-diameter steel tubes were bolted to the capping beam.

The sheet piles were left in the ground as a permanent part of the building, and the top slab of the car park served as bracing. Rebars were fixed to the sheet piles prior to pouring the slab of the car park, ensuring proper connection. The access ramp also has sheet pile retaining walls. The AZ 18 retaining wall has a total length of approximately 150 m.

In order to ensure the watertightness of the excavation and that of the future underground car park, several sealing measures were taken. The common AZ 18 double pile locks were welded together. In addition, the leading locks installed below the bottom slab were filled with a bituminous sealant .

After the installation of the steel sheet piles, excavation work began inside the retaining wall. To be able to work in the dry, the water table was lowered by two metres using pumps inside the excavation.

During the excavation operations, parts of the interlocks visible from the inside of the excavation were seal-welded.

To complete the car park waterproofing system, a small horizontal drainage system was created. Since the upper structure rests on concrete partitions, the connection between the bottom slab and the sheet piles did not have to take great loads.

As a result, 140 tonnes of steel sheet piles with a capping beam were installed within six weeks despite the complex shape of the excavation .

In the United States, during the construction of a mixed-use development consisting of 416 condominium residences with a large underground car park for 970 cars, geotechnical surveys revealed water at a depth of three metres below ground within the four-metre-thick layer of limestone. Beneath the limestone, there was a three-metre layer of loose sand, which lay over dense limestone and sandstone.

Instead of conventional temporary excavation support and an in-situ concrete wall, a bottom-up solution relying on sheet piles was chosen for the project as it provided for significant savings in construction costs and time. Being used as permanent retaining elements, sheet piles greatly simplified the construction process. First, a separate foundation system made of concrete bearing columns was constructed for seven seven-story houses. The sheet piles of the retaining wall for the 7.5-metre-deep excavation for the two-level underground car park were driven through the limestone until the design elevation was reached (Figure 6).

The construction was carried out using steel sheet piles as permanent retaining elements after the completion of excavation work and as temporary separation walls as the site was divided into several construction bays. Subdivision into several sections helped to reduce the concrete pour for the bottom slab, reaching manageable sizes.

About 3,000 metric tons of steel sheet piles were installed using a crane-mounted vibratory hammer.

About half of the piles remained in the ground at the end of construction as the permanent outer wall of the basement floor.

Nine-metre-long AZ 26 sheet piles with a web / flange thickness of 12.2 / 13.0 mm were selected to provide penetration through the hard soil to the design depth. A high-pressure pile equipped with a driving shoe was driven ahead of the sheet pile line to fragment the limestone, making it possible to drive the sheet piles without damage.

After pile driving was completed, formwork was installed at the heads of the sheet piles and a concrete cap was poured. Temporary ground anchors were incorporated in the capping beam that had been drilled into the sheet piles to provide temporary support until the 84 floor slabs were poured. A 1.2-metre-thick slab was then installed to provide the seal needed to start construction in the dry. The sheet piles, which started experiencing pressure as the water drained, were pressed against the concrete, sealing off water from below that was trying to seep through the interface between concrete and steel (17, 34-36).

The middle interlocks of the AZ 26 double piles had been welded before installation. After pumping out groundwater, the remaining interlocks were seal-welded. When the sealing was completed, the piles that served as permanent walls were cleaned and an aesthetic coating was applied (17, 18).

It looks to be more about case studies in France and USA and not related to the case study in St. Petersburg in which the data were used to as input to the present study .

Input data were collected from several underground car parks construction projects is St. Petersburg, Russia with complex geotechnical conditions especially in the city centre with high density of buildings many of which are of architectural and historical interest. Also, the city centre is characterized by weak clayey soils, and old river beds can be quite often found underground.

One of the most significant features of engineering-geological conditions of St. Petersburg, which must be taken into account in the construction of underground parking lots, is the distribution of weak clay soils. These sediments of lake-marine and lake-glacial genesis are capable of changing their behavior when the natural composition is disturbed. The impacts typical of a construction site in many cases lead to the disturbance of their natural structure. In this case, the soil ceases to work as a quasi-solid body and turns into a liquid-like medium. Therefore, it is necessary to develop technological solutions aimed at improving the reliability and safety of underground parking lots, as well as neighboring buildings and preserved architectural monuments.

Case-1

This case includes a residential development consisting of three buildings located within an area of 26,085 m^2 where two-level and one-level underground car parks were designed to be located at a depth of 8 m. During the construction, a pile foundation was driven. Loads per pile ranged from 120 to 150 tons, with loads on a pile group exceeding 300 tons. The piles were driven to a depth of 25 to 30 m (Figure 7).

The construction site is surrounded with four to five story 20th-century brick residential buildings. To the north and west. The buildings rest on strip rubble foundations, locally exhibit significant deformations in the form of oblique cracks, considered to be in critical condition with their foundations need to be reinforced to maintain their stability.

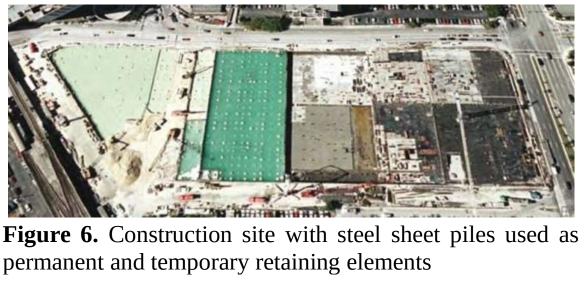


Figure 6. Construction site with steel sheet piles used as permanent and temporary retaining elements

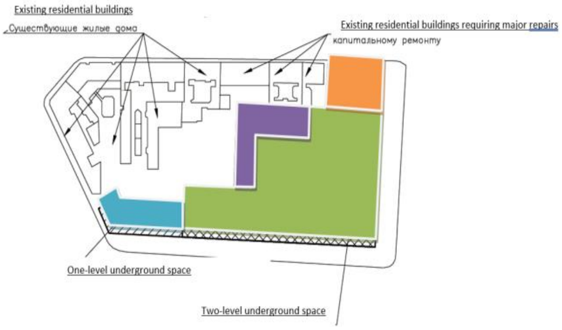


Figure 7. A building under construction with an underground car park in the central part of St. Petersburg

Moreover, the territory experienced demolishment of non-residential buildings of two to three story while remains of their concrete foundations still left on the ground. The territory can be used to implement reconstruction projects and build new structures provided they comply with height restriction laws and create frontage for the block.

The underground car park consists of five sections separated by contraction joints and is separated from the main part of the building by settlement joints.

When performing calculations of underground parking lots, it is necessary to take into account the engineering and geological conditions of the construction site. The most important indicators for calculations are the characteristics of soils at the construction site, which were laid in the basis of calculations of the presented study.

When constructing buildings and structures in dense urban areas, it is necessary to take measures to monitor the condition of buildings and structures located near the object under construction. To do this, terms of reference are drawn up, which should contain a rationale for monitoring operations, goals and objectives, parameters of the newly constructed building, parameters of existing buildings and structures within the area of influence, geotechnical conditions on site, and technical monitoring requirements (19-23).

Before starting the construction of new facilities, nearby buildings and structures must be inspected to check their external and internal state and the quality of their foundations. Particular attention should be paid to structures that were built several centuries ago and have not undergone major repairs.

Usually, there are several inspection stages :

-object monitoring (visual observation system, geodetic control);

-geological and hydrogeological monitoring,

-environmental and biological monitoring;
-analytical monitoring.

When construction works are carried out in a residential area, it is necessary to conduct geotechnical monitoring throughout the entire construction period of all objects located within a 30-metre radius from the construction site.

To prevent subsidence and soil compaction within the area of influence, soils under buildings and structures located in this area should be constantly monitored (24, 25). Also, geodetic monitoring needs to be performed in order to reveal irreversible processes in soils and predict settlement and tilt of buildings.

When erecting buildings and structures in complex geotechnical conditions, it is necessary to have a reliable forecast of the absolute and differential values of settlements. Therefore, in order to ensure the most accurate forecast, these values should be calculated for both foundations and superstructures.

For a newly constructed building, it is important to make sure that settlement in kept to a minimum. In the case being considered, this is due to the fact that the building is adjacent to other buildings, for which there is a settlement limit of 2 to 3 cm since they belong to Categories II and III in terms of their health.

In this geotechnical situation, it is advisable to consider a deep pile foundation solution for the projected development (26-28).

Geological surveys revealed the following soil types are presented in Table 1.

TABLE 1. Geological description of the construction site (up to a depth of 47.2 m)

N. п/п	Soil type	Designation
1.	Modern quaternary (QIV) deposits represented by man-made soils (tIV) (fill soils)	SL1
2.	Lacustrine and marine (lmIV) deposits (loose silty sands)	SL2
3.	Low-peat soils	SL3
4.	Silty sands (medium density)	SL4
5.	Soft sandy loams	SL5
6.	Fine and dense sands	SL6
7.	Very soft clay loams with interlayers of soft clay loam	SL7
8.	Upper quaternary (QIII) deposits of glaciolacustrine (lgIII) genesis represented by very soft clay loams with interlayers of soft clay loam; banded formations	SL8
9.	Soft clay loams, laminated	SL9
10.	Soft sandy loams, indistinct lamination	SL10
11.	Deposits of Glacial (gIII) genesis represented by hard sandy loams	SL11
12.	Silty sands (high density)	SL12
13.	Soft sandy loams with interlayers of hard sandy loams	SL13
14.	Middle Quaternary (QIII) sediments of glaciolacustrine (lgII) genesis represented by stiff clay loams with interlayers of semi-hard clay loams	SL14
15.	Deposits of Glacial (gII) genesis represented by hard sandy loams with soft interlayers	SL15

For a pile foundation, such soils as hard sandy loams of Glacial (gIII) genesis (SL11), as well as silty dense sands (SL12) lying at a depth of 19.8 to 24.9 m can be considered as a support layer for the piles (absolute elevation ranging from -14.7 to -22.0 m), as well as glaciolacustrine (lgII) soils ranging from stiff to hard loams (SL13, 14 and 15). However, soil layers SL11 and SL12 are not good options since their geology is irregular with inconsistent thickness values. As the upper layers are represented by moraine deposits, they have low bearing capacity, which also makes them unsuitable as support layers for piles. Based on the results of cone penetration tests, it is advisable to drive the piles down to layers SL13 and SL15 since these soils are located at a depth of more than 30 m from the surface.

Figure 8 (a, b) shows depth contour lines for the roofs of SL13 and SL15 layers, which are used as a support layer for piles.

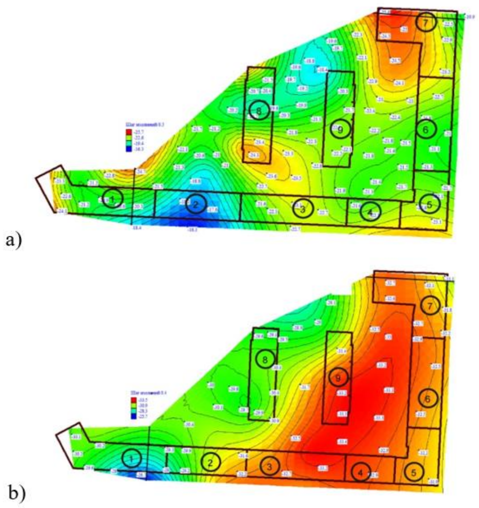


Figure 8. Depth contour lines for the roofs of the layers represented by: a) soft sandy loams with interlayers of hard sandy loams (gIII) (SL13) used as a support layer for piles in Sections 2, 3, 4, 5, 6, 8, 9; b) hard sandy loams with soft interlayers (SL15) and hard clay loams (SL16) used as a support layer for piles in Sections 1 and 7

The design depth (taking into account the depth of the capping beam) of the one-level underground car park will be about 5.2 m, and that of the two-level car park will be about 8.0 m.

When constructing a two-level underground car park, two excavation types can be used: top-down excavations and excavations with internal bracing.

In the case under consideration, it was decided to use a braced excavation with operations being performed in separate sections (29-32).

The sequence of operations in constructing the two-level underground car park is as follows: at the first step, a two-level metal bracing system is installed in the first trench, which serves as protection in constructing the underground car park; then, operations continue in a parallel trench; at the final step, the elements of the RC structures installed in the two parallel trenches are connected to form an underground structure in the following sequence (Figure 9):

-In each section, bracing struts are installed between the elements of two underground structures at a length of 3-4 m;

-The elements of the sheet piling are removed;

- Concrete is poured in the free space the two sections;

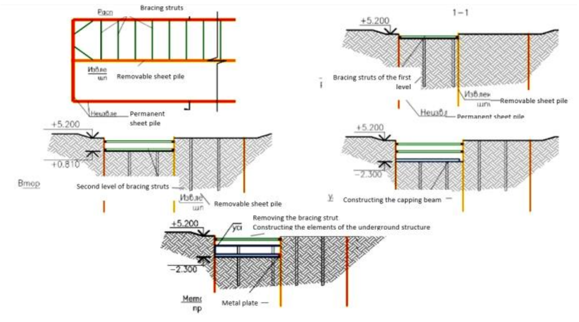
- The operations are repeated.

The advantages of using this method include the following:

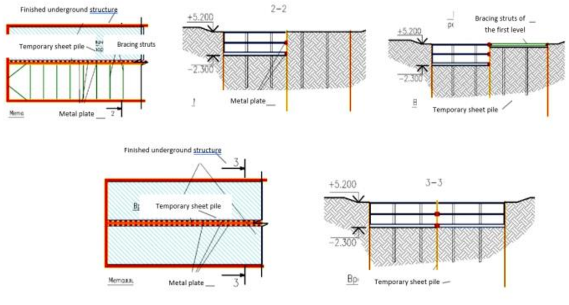
-excavation operations can be carried out in the required sequence and order;

-excavation costs are reduced;

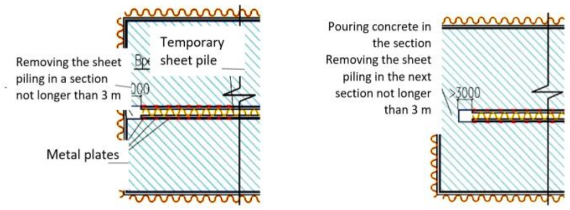
-it becomes possible to install expansion and contraction joints;



Step 1.



Step 2.



Step 3.

Figure 9. Steps in the construction of two-level underground facilities for separate sections using metal struts and temporary sheet piling

-there is no need to excavate the soil over the whole excavation area;

-if needed, it is possible to quickly suspend operations in one individual section .

The main disadvantages of this method include:

-the need for transporting and hauling a large number of sheet piles and installing a large number of bracing systems (33-35).

3. RESULTS AND DISCUSSION

To calculate the parameters of the sheet pile wall in the excavation being discussed, the Plaxis 2D software was used. This software calculates stress and strain in an array using the finite element method.

To minimize the impact of the two-level underground car park on existing buildings and structures, the following engineering solutions were developed:

-reinforcing the foundations and superstructures of nearby buildings in order to start work on the construction of the excavation;

-using permanent sheet piles to protect the excavation walls (sheet piles with a length of 24 metres);

-using a one-level or two-level bracing system made of pipes;

-testing piles before starting work to strengthen the foundations, main piles, and sheet piles, and digging a test excavation;

-monitoring the settlements of buildings and structures located in the area of influence, the level of vibration, and horizontal displacements of the sheet piles (17, 36).

The mathematical calculations were performed using 2D design schemes and the nonlinear Mora-Coulomb model, which made it possible to evaluate as accurately as possible the engineering-geological conditions in the construction area by analyzing the horizontal and vertical displacements of the sheet pile wall, the soil settlement at the bottom of the excavation, and the maximum settlement of the building located near the excavation.

The Coulomb-Mohr theory allowed us to describe mathematically the dependence of tangential stresses of a material on the magnitude of applied normal stresses, as well as to describe the dependence of ultimate tangential stresses on the average normal stress, since this theory is due to internal friction in a solid body. The Coulomb-Meer theory in the world practice is usually used to analyze the bearing capacity of soil masses. Under loading, soils work predominantly in shear along the surface with the lowest bearing capacity. Therefore, shear strength is the defining strength characteristic for soils. Failure is realized at the moment when the magnitude of shear (tangential) stress reaches the shear strength of the soil. Therefore, the relationship between normal stresses and tangential stresses is the strength criterion for soils.

Traditionally, analytical solutions based on the theory of limit equilibrium are used to calculate pit enclosures. In the considered case it is necessary to take into account the possibility of incomplete realization of active and passive pressures. For this purpose, it is necessary, one way or another, to set a nonlinear function of pressure dependence on wall displacement.

The simplest way of approximation of this function is the introduction of a bedding coefficient linking the values of pressure and wall displacement before reaching the active (when moving in the direction of pressure) or passive pressure (when moving in the opposite direction). To perform the calculation, special finite elements are used, whose deformability is determined by the value of the bed coefficient, and when the value of active or passive pressure is reached in the element, a nonlinear problem is solved by the method of initial stresses. In this connection, it is necessary to use models that take into account the difference of soil operation at the loading and unloading stages.

As observations of the deformations of envelope structures show, a characteristic feature of the operation of such structures is the long development of deformations in time. Calculation methods that do not take into account the time factor actually assume that after each excavation stage deformation occurs for a long period of time (until the "final" deformations are reached). In field conditions, excavation of an excavation pit may be performed relatively quickly, in which case the expected "final" deformations at all stages may not have time to develop. As a result, consideration of the time factor can have a significant impact on the nature of the performance of the envelope structure. Thus, as follows from this brief review of the theoretical assumptions of the calculation of pit envelope structures, this problem is a rather complex scientific and theoretical problem.

Based on the used Coulomb-Mohr model, the variables used in the calculations are the specific cohesion of the soil, which characterizes the resistance of the soil to shear (shear) in the absence of normal stresses at the shear (shear) site, soil characteristics, normal and tangential stresses.

Since the Plaxis 2D program was used to perform the calculations, which is a powerful and convenient finite element software package designed for two-dimensional calculations of deformations and stability of construction objects, the geometry was modeled with standard types of structural elements and loads using CAD drawing tools, which made it possible to quickly and efficiently create a finite element model. The Plaxis 2D program calculations are based on an ideal-elastic-plastic model with the Mohr-Coulomb strength criterion.

The Staged Construction mode allowed modeling the construction process by activating and deactivating soil clusters and structural elements at each stage of the calculation. Verified and stable iterative computation procedures were used. Using multi-threaded computing and a 64-bit computational kernel, Plaxis was able to compute this complex geotechnical model.

The Output program was used to output the data, which provides a variety of ways to display forces, displacements, stresses, and flows in the form of contours, vectors, and tables. Using the Curve manager option, graphs of different types of results at different stages of the calculation were created.

When calculating sheet pile parameters, 2D design diagrams were used. In calculations, the following three situations were considered: retaining wall at an elevation of +0.560 m; retaining wall at an elevation of -2.500 m; retaining wall in a complex section at an elevation of -2.500 m (adjacent building at a distance of 2.5 m).

The soil parameters used in the Plaxis 2D model when calculating the parameters of the retaining wall at an elevation of +0.560 m are presented in Table 2. The elevation of the surface is +6.000 m.

For the one-level car park, AZ 42-700N sheet piles are used. The cross-sectional area of each profile (S) is 259 cm^2/m, the moment of inertia (I) is 104,930 cm^3/m, the wall mass (m) is 203 kg/m^2, the elastic modulus (E) is 200,000 MPa, and the moment of resistance (W) is 4,205 cm^3/m. The bracing struts are made of 720x10 mm metal pipes located at an elevation of +4.700 m.

Based on the calculation results, it was found that the horizontal displacement of the sheet pile wall was 12 mm. It was also found that at the bottom of the excavation, there is an upward movement of soil by 166 mm, which is caused by the poor mechanical properties of the soils. The maximum settlement of the soil near the excavation was found to be 20 mm, which meets the requirements.

The soil parameters used in the Plaxis 2D model when calculating the parameters of the retaining wall at an elevation of -2.500 m are presented in Table 3. The elevation of the surface is +6.000 m.

For the two-level car park, HZ 1080M A-12 / AZ19-700N sheet piles are used. The cross-sectional area of each profile (S) is 324.8 cm^2/m, the moment of inertia (I) is 445,330 cm^3/m, the wall mass (m) is 270 kg/m^2, the

TABLE 2. Soil parameters for Problem 1

Designation	Elevation, m	Bulk density, kN/m^2	Volumetric frame weight, kN/m^2	Modulus of deformation, kN/m^2	Angle of internal friction, °	Poisson' s ration	Cohesion, kPa
SL1	+6.000	16.3	11.0	5,000	4	0.3	11
SL2	+4.500	17.8	13.2	5,000	23	0.3	1
SL3	+3.500	16.3	10.2	3,000	5	0.3	11.3
SL4	+0.500	19.7	15.4	15,000	28	0.3	3
SL5	-1.000	19.3	14.7	6,700	9	0.34	8
SL6	-3.000	20.7	17.0	38,000	35	0.3	3
SL7	-4.000	19.4	14.9	6,000	8	0.37	6
SL8	-10.000	18.2	1.29	5,000	3	0.37	6
SL9	-13.500	19.1	14.5	8,000	7	0.37	6
SL10	-17.500	19.9	15.8	11,000	10	0.34	31
SL11	-18.000	22.1	19.0	21,000	22	0.34	6
SL12	-24.500	21.0	17.5	34,000	33	0.3	7
SL13	-28.500	21.8	18.8	18,000	20	0.34	25

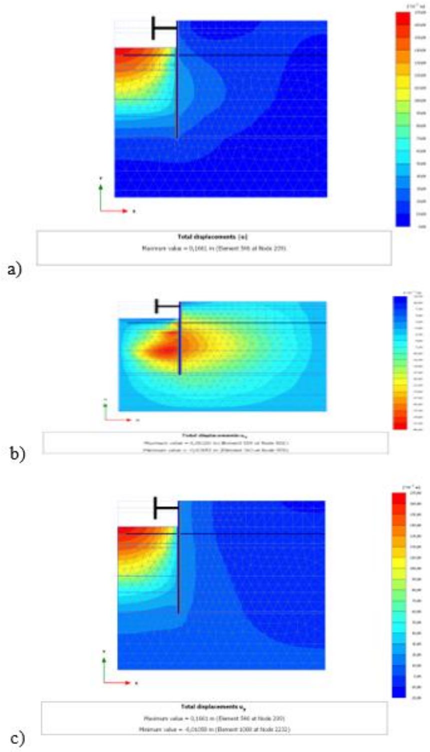


Figure 10. Calculation results for Problem 1: a) total displacements; b) total displacements along the x-axis; c) total displacements along the y-axis

TABLE 3. Soil parameters for Problem 2

Designation	Elevation, m	Bulk density, kN/m^2	Volumetric frame weight, kN/m^2	Modulus of deformation, kN/m^2	Angle of internal friction, °	Poisson' s ration	Cohesion, kPa
SL1	+6.000	16.3	11.0	5,000	4	0.3	11
SL2	+4.500	17.8	13.2	5,000	23	0.3	1
SL3	+3.000	16.3	10.2	3,000	5	0.3	11.3
SL4	+0.500	19.7	15.4	15,000	28	0.3	3
SL5	-1.000	19.3	14.7	6,700	9	0.34	8
SL6	-2.000	20.7	17.0	38,000	35	0.3	3
SL7	-3.000	19.4	14.9	6,000	8	0.37	6
SL8	-6.000	18.2	1.29	5,000	3	0.37	6
SL9	-9.000	19.1	14.5	8,000	7	0.37	6
SL10	-15.000	19.9	15.8	11,000	10	0.34	31
SL11	-18.000	22.1	19.0	21,000	22	0.34	6
SL12	-23.000	21.0	17.5	34,000	33	0.3	7
SL13	-30.000	21.8	18.8	18,000	20	0.34	25

elastic modulus (E) is 200,000 MPa, and the moment of resistance (W) is 7,790 cm^3/m.

The bracing struts are made of 720x10 mm metal pipes, with the first level located at an elevation of +4.700 m, and the second one at an elevation of +2.250 m.

Based on the calculation results, it was found that the horizontal displacement of the sheet pile wall was 20 mm. At the bottom of the excavation, there is an upward movement of soil by 292 mm, which is caused by the poor mechanical properties of the soils. The maximum settlement of the soil near the excavation was found to be 20 mm, which meets the requirements.

The soil parameters used in the Plaxis 2D model when calculating the parameters of the retaining wall at an elevation of -2.500 m (with an adjacent building at a distance of 2.5 m) are presented in Table 4. The elevation of the surface is +6.000 m.

In this case, a combination of AU-14 sheet piles and 820x10 mm pipe piles reinforced with 40k5 I-beams and filled with B20 concrete is used.

Sheet pile parameters in this case are as follows: the cross-sectional area (S) is 132 cm^2/m, the wall mass (m) is 104 kg/m^2, the moment of inertia (I) is 28,680 cm^3/m, the moment of resistance (W) is 1,405 cm^3/m, and the elastic modulus (E) is 200,000 MPa.

Composite pile parameters are as follows: the cross-sectional area (S) is 5,278 cm^2/m, the wall mass (m) is 1,248 kg/m^2, the axial stiffness (EA) is 20,790,262

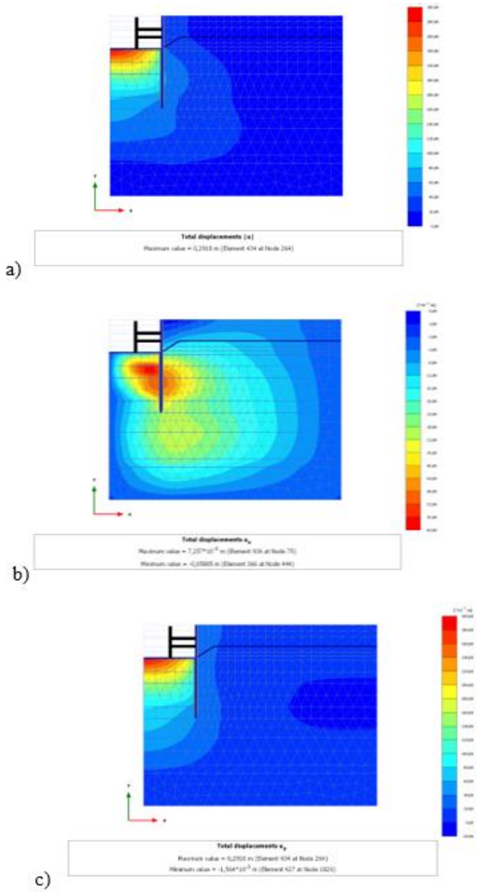


Figure 11. Calculation results for Problem 2: a) total displacements; b) total displacements along the x-axis; c) total displacements along the y-axis

TABLE 4. Soil parameters for Problem 3

Designation	Elevation, m	Bulk density, kN/m^2	Volumetric frame weight, kN/m^2	Modulus of deformation, kN/m^2	Angle of internal friction, °	Poisson’s ration	Cohesion, kPa
SL1	+6.000	16.3	11.0	5,000	4	0.3	11
SL2	+4.500	17.8	13.2	5,000	23	0.3	1
SL3	+3.000	16.3	10.2	3,000	5	0.3	11.3
SL4	+0.500	19.7	15.4	15,000	28	0.3	3
SL5	-1.000	19.3	14.7	6,700	9	0.34	8
SL6	-2.000	20.7	17.0	38,000	35	0.3	3
SL7	-3.000	19.4	14.9	6,000	8	0.37	6
SL8	-6.000	18.2	1.29	5,000	3	0.37	6
SL9	-9.000	19.1	14.5	8,000	7	0.37	6
SL10	-15.000	19.9	15.8	11,000	10	0.34	31
SL11	-18.000	22.1	19.0	21,000	22	0.34	6
SL12	-23.000	21.0	17.5	34,000	33	0.3	7
SL13	-30.000	21.8	18.8	18,000	20	0.34	25

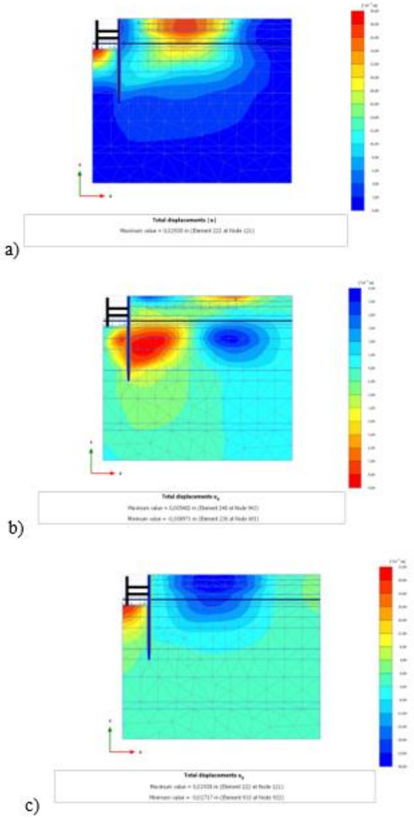


Figure 12. Calculation results for Problem 3: a) total displacements; b) total displacements along the x-axis; c) total displacements along the y-axis

kN/m^2, and the bending stiffness (EI) is 1,046,064 kN/m^2.

The bracing struts are made of 720x10 mm metal pipes, with the first level located at an elevation of +4.700 m, and the second one at an elevation of +2.250 m.

In Plaxis 2D, two slabs located at a minimum distance from each other were used in solving this problem. Between the slabs, the space is filled with elastic soil, which only transfers the load from one slab to another.

Based on the calculation results, it was found that the horizontal displacement of the sheet pile wall was 6 mm. At the bottom of the excavation, there is an upward movement of soil by 32 mm, which is caused by the poor mechanical properties of the soils. The maximum settlement of the building located at a distance of 2 metres from the excavation was found to be 30 mm, which meets the requirements.

Poor mechanical properties of soils will cause changes in the initial stress state, which will lead to a loss of stability of the soil layer adjacent to the foundation. This will especially affect the part of the foundation that is directly adjacent to the excavation. What is the most dangerous is that in the process of excavation, plastic soil deformations may develop, with the foundation footing of the existing building being squeezed out towards the excavation. Therefore, when constructing an excavation near existing buildings and structures, it is very important to take into account the relative positions of the foundation footings and the excavation being constructed.

It is also very important to understand that the slightest mistake in the design or implementation stage of a project for building a new underground car park or renovating buildings in difficult geotechnical conditions can lead to the destruction of objects of architectural interest that have survived to this day, as well as to the resettlement of the surrounding residential buildings.

It can be seen from the earlier researches of the authors of the article that numerical analysis allows to determine quite accurately the parameters of interest, namely, to predict the development of settlements due to improved soil models taking into account its nonlinear operation under the action of load, to determine the stress-strain state of the system "soil mass - sheet pile wall - surrounding building", and also that when designing underground engineering facilities in complex engineering-geological conditions it is best to use for mathematical calculations the following parameters: «soil mass - sheet pile wall - surrounding building».

Objectivity and correctness of the authors' judgments are confirmed by the use of proven methods of mathematical modeling, implemented in the engineering software package Plaxis 2D, as well as positive results of experimental studies in full-scale conditions. The results of this study are confirmed by the fact that the calculated values coincide with the experimental data by 92%. According to the results of the calculations, it was found that the horizontal displacements of the sheet pile wall in the 1st, 2nd and 3rd problem were: 12 mm, 20 mm and 6 mm, respectively. Whereas in-situ measurements of sheet pile wall displacement showed values of 11, 19 and 5 mm. The same convergence was shown by the calculated soil displacement values and calculations of the maximum settlement of the soil near the excavation. This confirms the reliability of the calculations.

The innovativeness of the performed research consists in the use of new approaches and solutions for the construction of underground parking lots and multi-level parking lots in complex engineering and geological conditions of megacities.

4. CONCLUSIONS

To ensure the safety of objects of architectural interest during the construction of underground parking lots and the safety of nearby buildings that may be affected by construction or renovation works, it is necessary to conduct comprehensive geotechnical monitoring. Due to the complexity and multifactor nature of geotechnical problems, it is possible to find a solution in most cases only using modern numerical methods based on physical and geomechanical nonlinear models.

Objectivity and correctness of the authors' judgments are confirmed by the application of proven methods of mathematical modeling, implemented in the engineering software package Plaxis 2D, as well as positive results of experimental and field studies.

The results of this study are confirmed by the fact that the calculated values coincide with the experimental data by 92%. According to the results of the calculations, it was found that the horizontal displacement of the spindle wall at setting 1, 2 and 3 tasks amounted to: 12 mm, 20 mm and 6 mm, respectively. At the same time, field measurements of the spindle wall displacement showed values of 11, 19 and 5 mm. The same convergence was shown by the ground motion calculation and the calculation of the maximum settlement of the ground near the pit. This confirms the accuracy of the calculations.

The results presented in the paper are of great scientific and practical importance for improving the reliability and safety of preserved architectural monuments during the construction of underground parking lots, as well as the safety of neighboring buildings that may be affected by construction or reconstruction.

The results of tests and calculations can be used by companies engaged in the design and construction of underground parking lots in complex engineering and geological conditions, as well as by organizations engaged in the control of serviceability of structures, design of the zero cycle of foundations and earthworks. The results can be used by scientific-technical and research institutes in the implementation of similar projects, as well as in the educational process of training specialists of the highest category.

5. ACKNOWLEDGMENT

The authors of the article express their gratitude to their colleagues from the Department of Construction of Mining Enterprises and Underground Structures, as well as to the St. Petersburg Mining University.

6. REFERENCES

1. Alliluev V. Layout and arrangement of parking houses in development of underground space in unbuilt city areas. Mining Informational and Analytical Bulletin. 2018(4):218-25. 10.25018/0236-1493-2018-4-0-218-225

2. Kushwaha A, Saini A, Shukla S. Analysis of Parking Facility in an Urban City Using Geospatial Technology: A Case Study Mahanagar Underground Parking Lucknow. International Journal for Research in Applied Science & Engineering Technology,. 2022;10(11). 10.22214/ijraset.2022.40375
3. Zhang P, Chen Z, Liu H. Study on the layout method of Urban underground parking system-a case of underground parking system in the Central business District in linping New City of Hangzhou. Sustainable Cities and Society. 2019;46:101404. 10.1016/j.scs.2018.12.032
4. Martyushev N, Avramchuk V, Faerman V, editors. International Conference on Information Technologies in Business and Industry 2016. Journal of Physics: Conference Series (см в книгах); 2017: Institute of Physics and IOP Publishing Limited. 10.1088/1742-6596/803/1/011001
5. Shutaleva A, Nikonova Z, Savchenko I, Martyushev N. Environmental education for sustainable development in Russia. Sustainability. 2020;12(18):7742. 10.3390/su12187742
6. Visweswara Rao S. SCK, Srihari J., Ruthik Reddy D., . Ventilation system for a multi level underground parking space. International Journal of Scientific Research in Engineering and Management. 2022;6(1):1-5. 10.55041/IJSREM11465
7. Goel RK, Singh B, Zhao J. Underground infrastructures: planning, design, and construction: butterworth-heinemann; 2012.
8. Komolov V, Belikov A, Kankhva V, Mezina N, editors. Assessment of the impact of the construction of semi-buried structures on the surrounding buildings and the road system. IOP Conference Series: Materials Science and Engineering; 2020: IOP Publishing. 10.1088/1757-899X/918/1/012027
9. Iraji A. Reinforced soil wall analysis under working stress conditions using a two phase model with the introduction of a new design parameter. International Journal of Engineering, Transactions C: Aspects. 2019;32(12):1762-72. 10.5829/IJE.2019.32.12C.09
10. Chandiwala A, Vasanwala S. Experimental Study of Lateral Loading on Piled Raft Foundations on Sandy Soil. International Journal of Engineering, Transactions A: Basics. 2023;36(1):28-34. 10.5829/ije.2023.36.01a.04
11. Mulyawati IB, Riza M, Dermawan H, Pratiwi V. Numerical simulation of embankment settlement in vacuum preloading systems. International Journal of Engineering, , Transactions A: Basics. 2023;36(4):817-23. 10.5829/ije.2023.36.04a.18
12. Long M. Database for retaining wall and ground movements due to deep excavations. Journal of Geotechnical and Geoenvironmental Engineering. 2001;127(3):203-24. 10.1061/(ASCE)1090-0241(2001)127:3(203)
13. Long M, Menkiti C, Skipper J, Brangan C, Looby M. Retaining wall behaviour in Dublin's estuarine deposits, Ireland. Proceedings of the Institution of Civil Engineers-Geotechnical Engineering. 2012;165(6):351-65. 10.1680/geng.10.00037
14. Enders A, Kuhlmann U. Design of semi-compact Z-shaped steel sheet pile walls. Modern Trends in Research on Steel, Aluminium and Composite Structures: Routledge; 2021. p. 164-70.
15. Wang J, Xiang H, Yan J. Numerical simulation of steel sheet pile support structures in foundation pit excavation. International Journal of Geomechanics. 2019;19(4):05019002. 10.1061/(ASCE)GM.1943-5622.0001373
16. Yang Y, Liu W, Hu A, Yang Z, Shen D, editors. Study on Larsen Steel Sheet Pile as the Deep Base Pit Support. IOP Conference Series: Earth and Environmental Science; 2020: IOP Publishing. 10.1088/1755-1315/525/1/012017
17. Ou C-Y. Deep excavation: Theory and practice: Crc Press; 2014.
18. Lezhava I, Golubev G. Problems of Construction Activity in the Underground Space of Downtown Moscow. Soil Mechanics and Foundation Engineering. 2004;41:119-24. 10.1023/B:SMAF.0000046043.55484.8e
19. Ivanov VV, Sidorenko SA, Sidorenko AA. Justification of the method for determination the optimum performance of limestone quarry for steel and cement production. Biosciences Biotechnology Research Asia. 2015;12(2):1797-803. 10.13005/bbra/1844
20. Braila N, Znobishcheva M, Panchenko N, Kostyshak M, editors. Prospects of underground parking during the reconstruction in the center of St. Petersburg. E3S Web of Conferences; 2020: EDP Sciences. 10.1051/e3sconf/202016404008
21. Trushko OV, Trushko VL, Demenkov PA. Arrangement of multistory underground parking garages in complex engineering and geological environment. International Journal of Mathematical, Engineering and Management Sciences. 2020;5(5):897. 10.33889/IJMEMS.2020.5.5.069
22. Dashko RE, Lange IY. Engineering-geological aspects of negative consequences of contamination of dispersive soils by petroleum products. Записки Горного института. 2017;228:624-30. 10.25515/PMI.20176.624
23. Dashko R, Kotiukov P. Analysis of construction accident in Saint Petersburg based on consideration of underground space as a contaminated multicomponent system. International Multidisciplinary Scientific GeoConference: SGEM. 2017;17:67-74. 10.5593/sgem2017/51/S20.010
24. Granata R, Gioia F, Cotecchia V. Construction of a large underground parking in Bari downtown. Geotechnical Aspects of Underground Construction in Soft Ground: CRC Press; 2012. p. 585-94.
25. Mangushev R, Nikiforova N. Technological Settlements of the Surrounding Buildings during the Construction of Deep Pit Fences. Soil Mechanics and Foundation Engineering. 2023;60(1):15-21. 10.1007/s11204-023-09858-3
26. Kulikova EY, Konyukhov D. Accident risk monitoring in underground space development. MIAB Mining Inf Anal Bull. 2022(1):97-103. 10.25018/0236_1493_2022_1_0_97
27. Mangushev R, Osokin A, Levinskaya P. Prospects for the construction of underground parkings under conditions of tight development of the historic center of St. Petersburg. Housing Construction. 2019;4:3-18. 10.31659/0044-4472-2019-4-3-18
28. Ilyichev V, Nikiforova N, Konnov A, editors. Technological Soil Mechanics in Underground Construction. IOP Conference Series: Materials Science and Engineering; 2021: IOP Publishing. 10.31659/0044-4472-2019-4-3-18
29. Mangushev R, Osokin A, Diakonov I, Kalach F, editors. Constructive and technological solutions for underground space safety amidst dense historical buildings and weak foundation soils. E3S Web of Conferences; 2023: EDP Sciences. 10.1051/e3sconf/202337102002
30. Belyakov N, Smirnova O, Alekseev A, Tan H. Numerical simulation of the mechanical behavior of fiber-reinforced cement composites subjected dynamic loading. Applied Sciences. 2021;11(3):1112. 10.3390/app11031112
31. Lebedev M, Karasev M, Belyakov N, Basova L. Face Stability in Heavy Clay: Theory and Practice. Journal of Mining Science. 2022;58(2):234-45. 10.1134/S1062739122020077
32. Il'Ichev V, Mangushev R, Nikiforova N. Development of underground space in large Russian cities. Soil mechanics and foundation engineering. 2012;49(2):63-7. 10.1007/s11204-012-9168-6
33. Pushkarev A, Sokolov N, Mihailov A. Methods and technology of reinforcing a leaning slope. Transportation Research Procedia. 2021;57:538-46. 10.1016/j.trpro.2021.09.082
34. Shashkin A, Shashkin K, Dashko R, editors. Analysis of causes of deformations in historic buildings on weak clay soils. Geotechnics Fundamentals and Applications in Construction:

New Materials, Structures, Technologies and Calculations–Proceedings of the International Conference on Geotechnics Fundamentals and Applications in Construction: New Materials, Structures, Technologies and Calculations, GFAC; 2019. 10.1201/9780429058882-64

35. Shashkin A, Shashkin K, Ulitsky V, editors. Calculation of soil–transport structure interaction. Transportation Soil Engineering in Cold Regions, Volume 2: Proceedings of TRANSOILCOLD 2019; 2020: Springer. 10.1007/978-981-15-0454-9_14

36. Evseev N, Shashkin K, Vasenin V, Shashkin V. Deformation criteria for ensuring mechanical safety of historical brick buildings during reconstruction. Reconstruction and Restoration of Architectural Heritage 2021: CRC Press; 2021. p. 204-9.

Persian Abstract

چکیده

در کلان شهرهای مدرن، هر روز مردم با مشکل یافتن مکانی برای پارک خودروها مواجه می شوند، این مشکل به ویژه در مرکز شهر حادتر است .با توجه به رشد سریع تعداد خودروها، تعداد فضاهای پارک موجود به شدت کم است .بنابراین راه حل بهینه برای ساخت و ساز و بازسازی ساختمان ها در کلانشهرهای بزرگ، استفاده منطقی از فضای زیرزمینی، یعنی احداث پارکینگ های زیرزمینی است .کار حاضر تحلیلی بر ساخت پارکینگ های زیرزمینی و چند طبقه در شهرهای بزرگ از طریق پرونده ای در شهر سن پترزبورگ روسیه است .راه حل های بهینه برای ساخت پارکینگ های زیرزمینی در شرایط پیچیده ژئوتکنیکی با استفاده از برنامه Plaxis 2D در نظر گرفته است .هدف از این مطالعه توسعه راه حل های فناورانه برای ساخت پارکینگ های زیرزمینی در شرایط پیچیده مهندسی و زمین شناسی کلانشهرهای بزرگ با استفاده از حصار شمع ورق می باشد . داده های ورودی از مجموعه داده های مهندسی و بررسی های زمین شناسی موجود به دست آمده از سایت های ساخت و ساز جمع آوری شد .روش مورد استفاده، نمودارهای طراحی دو بعدی و یک مدل غیر خطی موهر-کلمب بود که امکان ارزیابی دقیق ترین شرایط ژئوتکنیکی در منطقه ساخت و ساز را با تجزیه و تحلیل جابجایی های افقی و عمودی دیوار شمع ورق، نشست خاک در منطقه فراهم کرد .پایین حفاری، و حداکثر استقرار یک ساختمان واقع در نزدیکی حفاری .در نتیجه، توصیه‌های حرفه‌ای برای ساخت پارکینگ‌های زیرزمینی و چند طبقه در شرایط پیچیده ژئوتکنیکی ارائه شد: لازم است پشتیبانی ژئوتکنیکی پیچیده ساخت و ساز انجام شود .هنگام ساخت یک گودال در شرایط سخت مهندسی و زمین شناسی، یک گودال با سیستم های فلزی فاصله ساز (چال باز) بسازید .هنگام ساخت گودال در نزدیکی ساختمان ها و سازه های موجود، توجه به موقعیت نسبی پایه فونداسیون و گودال در حال ساخت بسیار مهم است .انجام محاسبات دقیق و مطمئن در طراحی و اجرای ساخت پارکینگ زیرزمینی به عنوان بخشی از ساخت و ساز جدید یا بازسازی ساختمان های قبلی ساخته شده در شرایط سخت مهندسی و زمین شناسی .نویسندگان مقاله راه حل های فن آوری جدیدی را ایجاد کرده اند که از اهمیت علمی و عملی زیادی برای بهبود قابلیت اطمینان و ایمنی بناهای معماری حفظ شده در طول ساخت پارکینگ های زیرزمینی و همچنین ایمنی ساختمان های مجاور که ممکن است تحت تأثیر ساخت و ساز قرار گیرند، ایجاد کرده اند.

IJE TRANSACTIONS B: Applications Vol. 37 No. 02, (February 2024) 237-251

International Journal of Engineering

Journal Homepage: www.ije.ir

A Novel Design and Simulation of a Nano Prosthetic Artificial Heart Valves

S. Mahmood Ali*

Biomedical Engineering Department, University of Technology, Iraq

PAPER INFO

Paper history:
Received 14 April 2023
Received in revised form 01 November 2023
Accepted 02 November 2023

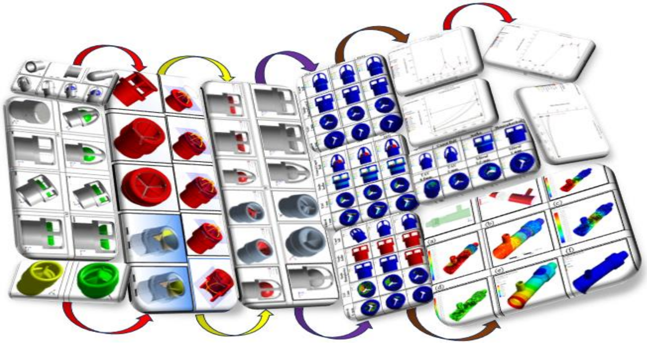
Keywords:
Artificial Caged Ball Valve
Single-Leaflet Heart Valves
Heart Valves Simulation
Tricuspid Aortic Valve
PSN4
SIBSTAR103T

ABSTRACT

Heart valve replacement is a major health burden and is required by millions of people worldwide, which invites the continuous need to discover and manufacture more effective and permanent artificial replacements. In the present work, unique models of eight artificial heart valves were designed and examined using seven synthetic and nanocomposite materials. The designed valves were examined to determine the best designs and materials in terms of durability, flexibility, and energy consumption, and to improve the biomechanical performance by using the Response Surface Methodology (RSM) and the Design Expert System 13. The highest values of the equivalent stress due to the applied blood pressure on the moving parts on each type of manufactured heart valve occur in valves with three dimensions moving parts, reached in the mitral tri-leaflet valve 14.13 MPa, followed by the tricuspid aortic valve. The equivalent stresses for other types of valves produced with simple surface action were lower than 2 MPa. The strain energy that is expended during the process of diastole and systole was found to be directly proportional to the strength and flexibility of the materials used. The energy consumption rates decrease when using highly elastic materials such as TPE and PSN4. The values of this energy also increase with an increase in the area of the moving parts of the valve, especially when faced with the process of closing blood flow, as with the use of the tricuspid aortic valve (TAV). The highest total deformation resulted in the valve body when using TPU, followed by TPE, nylon, PETG, and PLA, while the lowest deformation rates were observed when using PSN4, which ranged from 5x10^5 to 0.1 mm, followed by SIBSTAR103 nanostructured rubber. The obtained values of stress safety factors were decreased with the complexity of the movement for the moving parts of the valve. The highest rates were recorded when using the tricuspid mitral valve, reaching 2.45 when using the high-strength and flexible PSN4 nanomaterial. It can be concluded that the best materials for manufacturing these four types of valves are the PSN4, followed by SIBSTAR103T, TPU, and TPE. The use of PETG, PLA, and nylon materials is not recommended for the manufacture of any prosthetic heart valves, due to their lack of strength, flexibility, and high brittleness, especially for PETG and PLA materials. It was also noted here that PSN4 is the only material suitable for the manufacture of mitral tri-leaflet and tricuspid mitral valve artificial valves. For other types of valves manufactured with a single leaflet, high safety stress factors were obtained because their movement is simple, flat, and in one direction, where the highest values were observed when testing a single hemispherical leaflet type valve, then the conical caged ball and the caged ball type, respectively.

doi: *10.5829/ije.2024.37.02b.03*

Graphical Abstract



---

*Corresponding author email: *30249@uotechnology.edu.iq* (S. Mahmood Ali)

Please cite this article as: Mahmood Ali S. A Novel Design and Simulation of a Nano Prosthetic Artificial Heart Valves. International Journal of Engineering, Transactions B: Applications. 2024;37(02):237-51.

NOMENCLATURE

CFD	Computational Fluid Dynamics	a	Dimensionless parameters; $a = \delta/R0$
FVM	Finite volume method	b	Dimensionless parameters; $b = (m/L_0)^2$
PETG	Polyethylene Terephthalate Glycol	k	The flow consistency index
PLA	Polylactic Acid	L	The half-length of an artery model (cm)
PSN4	Polyetherimide/Silicone Rubber	L_0	half-length of the artery stenosis(cm)
RSM	Response Surface Methodology	m	Constant value
SIBSTAR103T	Isobutylene-B-Styrene Rubber Triblock	n	The power-law index
TAV	Tricuspid Aortic Valve	p	The blood pressure (mmHg)
TPE	Thermoplastic Polyurethane Elastomer	Q	Blood flow rate (liter/min.)
TPU	Thermoplastic Polyurethane	Q_0	Initial value of blood flow rate (cm^3/sec)
VHD	Valvular Heart Disease	r	Radial coordinate transformation, $r = yR$,
τ_{rz}, τ_{rr}, and τ_{zz}	The shear stresses of the model	$R(z)$	The a stenosed artery's radius (mm)
τ_{ij}	The Reynolds stress	R_0	The radius of a normal artery (cm)
ρ	Blood density (kg/m^3)	Re	Reynolds number
λ	The resistance to the blood flow	S	Modulus of the mean rate-of-strain
μ	Dynamic viscosity of blood (Pa s)	S_0	Initial blood vessel cross-sectional area (cm^3)
δ_{ij}	Parameter controls the percentage stenosis.	V	The bulk stream-wise velocity related Re
δ	The highest depth of the stenosis (cm)	Vr	The blood radial velocity
v	The blood parabolic velocity profile (m/s)		

1. INTRODUCTION

Congenital and rheumatic valvular heart disease (VHD) is the most important cause of heart failure around the world (1-3), which can cause arrhythmias, stroke, heart failure, and potentially life-threatening complications and need continuous clinical attention. More than 30 million people suffer from it. In the United States alone, approximately 90,000 heart valve replacements and more than 300,000 worldwide are performed annually with biomechanical or prosthetic valves (4).

Each human heart has four valves: the mitral, the tricuspid, the aortic, and the pulmonary valve (1, 5). There are various forms of aortic valve disease, namely aortic regurgitation, aortic stenosis, congenital aortic valve disease, and atrial septal defect (6). Artificial heart valves are used to replace heart valves that have been damaged by disease or wear and tear for time duration (7).

Mechanical heart valves are made of materials of much higher strength than those found in human soft tissues. There are basic types of mechanical heart valve designs including, the gaged ball, single leaflet, bicuspid, and tricuspid valves (8). The main advantages of the mechanical valve are their durability, and they are well-suited for patients of all ages. These Artificial heart valves rarely need replacement and often last for the rest of the patient's life. The main drawback of a mechanical valve is its tendency to form blood clots on its surfaces, so patients need lifelong blood-thinning medications and lifestyle modifications (9).

The use of prosthetic artificial heart valves and continuous improvements in valve designs have reduced the death rate of heart patients (9). The selection of a prosthetic valve that is appropriate in shape and size is a key factor for the success of surgery (10). However, the ability of the valve geometry to provide adequate mechanical function immediately after implantation is critical to patient survival. Therefore, the successful application of tissue engineering in the development of heart valves will require that they maintain similar mechanical properties to that of the original valves, and the ability to repair, remodel, and regenerate heart tissue (11, 12). So far, no satisfactory mechanism has been found to achieve a dynamically stable and durable implantable heart valve (13). The latest designs of mechanical heart valves provide relatively superior hemodynamics, but high-shear platelets and blood cells still pose significant challenges (14).

Composite materials are one of the most well-known materials for various functions that reduce the structure's weight without reducing its strength. Polymer composites exploit a wide range of applications due to their excellent performance in mechanical, electrical, and thermal properties (such as medical, electronic, automobile, and aerospace) (15).

The current work aims to design and simulate a unique model of eight artificial heart valves using seven synthetic and nanocomposite materials. The designed valves will be examined to determine the best designs and materials in terms of durability, flexibility, and energy consumption performances by using the Response Surface Methodology (RSM) and Expert Design 13 system.

2. MATERIALS METHODOLOGY

2. 1. Design of a Human Artificial Heart Valves

The mechanical prosthetic heart valves are an effective engineering device used to replace the damaged or diseased natural valves of the heart and are characterized by long performance and high reliability, but they may lead to possible complications such as hemolysis and thrombosis (16). Durability is the most important factor in choosing the right materials for polymeric heart valves to ensure strong mechanical properties that give the valve low creep, high rigidity with self-healing, and the ability to maintain structural integrity at repeated periodic loading and unloading of blood to prevent life-long blood clotting, and it is less prone to calcification and failure compared to bio-prosthetic valves (5).

The most important factors that must be considered when choosing a valve are the patient's age, comorbidities, and life expectancy (8). The process of designing an artificial heart valve needs to evaluate its mechanical performance and blood flow pattern and study both leaflet deformation/stress. As the leaflet is subjected to fluctuating stress, the large lifetime of the stress requires flexibility to open and close the blood flow path as a major function of the heart valve (17). The most important design requirements for valves are those that cause minimal shock to the elements of the structure surrounding the valve, blood vessels, and endothelial tissues of the heart, with good resistance to mechanical and structural abrasion. It works to reduce the chances of deposition of platelets and thrombus, to be non-dissolving materials, to be stable and not absorb blood components, and to be sterilizable with high surface quality (18).

In this work, eight artificial heart valves were designed and analyzed. Table 1 lists the main physical and mechanical properties of the designed materials.

Figure 1 illustrates the designing steps for the artificial caged ball heart prosthetic valve. The caged ball valve is the first artificial heart valve designed and implanted in the human body. The valve was initially made of stainless steel (12). The introduction of granulated ball valves is a major advance in the treatment of patients.

Now, there are several thousand patients who have such valves implanted (8, 19). The caged ball valve design is still the valve of choice for some surgeons (18, 20). It consists of a celestial ball with a circular stitching ring and a cage of three or four plastic or metal brackets (usually of titanium rods) that determines the path of movement of the ball (21).

The ball blocks the valve opening to prevent the reverse flow of blood. Its diameter is greater than the inner diameter of a circular ring.

TABLE 1. The main physical and mechanical properties of the designed materials (12, 22-29)

Filament materials	Density (g/cm^3)	Nozzle temp. (°C)	Heated bed temp. (°C)	Printing speed (mm/s)	Tensile Strength (MPa)	Tensile Elongation (%)	Impact Strength (kJ/m^2)	Flexural Strength (MPa)	Flexural Modulus (GPa)
Nylon (White)	1.16	240	85	35	6.37	231.30	28.30	8.34	0.20
PETG (Blue)	1.27	240	75	50	4.90	120.00	8.70	70.61	2.15
(TPU) Transparent (purple, orange, and yellow)	1.21	235	60	40 – 60	16.80	550.00	34.40	1.80	0.08
TPE (Red)	1.14	235	30	20-40	14.00	990.00	-	80.00	1.00
PLA (White and, Transparent purple)	1.24	200	40	50	6.08	4.40	4.20	6.47	2.75
SIBSTAR 103T	0.92	-	-	-	18.10	506.00	-	-	5.92
PSN4	1.27	-	--	-	39.13	108.50	102.60	73.24	2.43

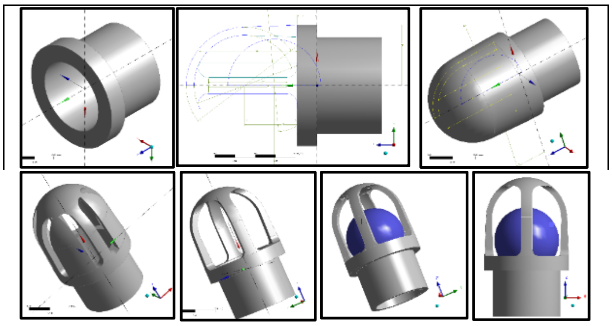
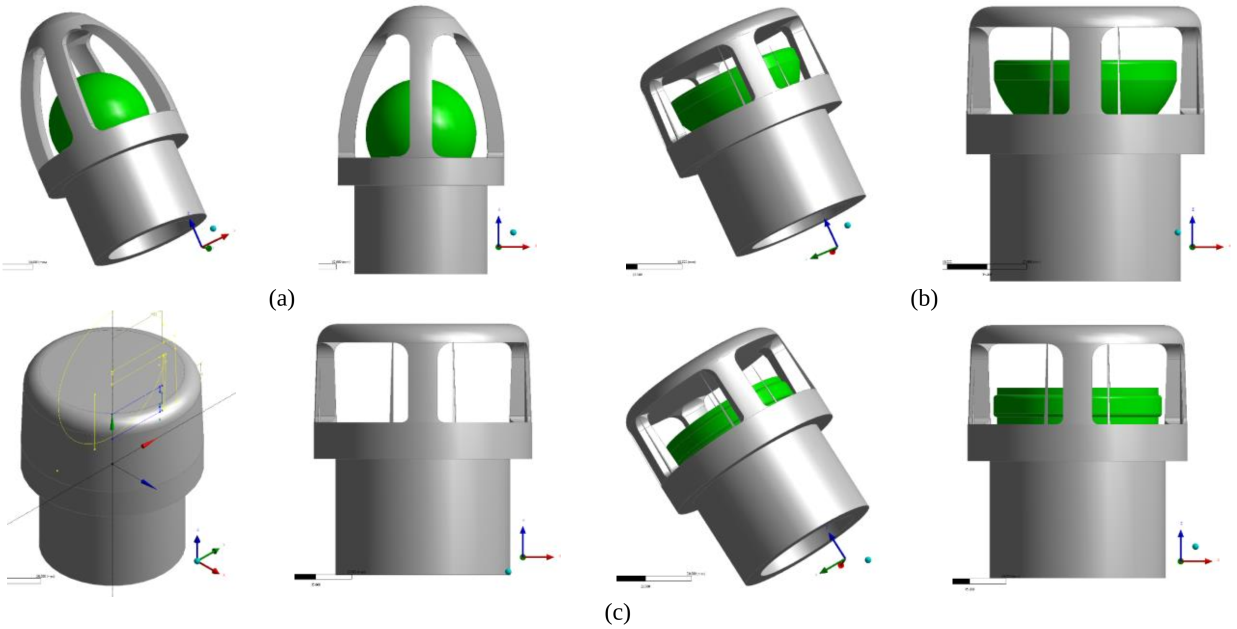


Figure 1. The designed steps of the artificial caged ball mechanical heart valve

These valves were designed to allow blood to flow in only one direction, and act to cause blood to flow from the heart to the aorta when the pressure of the heart exerts on the blood (and the valve ball) exceeds the pressure in the aorta. After the blood leaves the heart, the pressure inside the heart drops drastically, then the negative pressure sucks the ball of the valve back, which will prevent the blood from returning to the heart.

Figure 2a shows the artificial conical caged ball heart valve. This valve works with the same principle as the previous valve, except that the cage that holds the ball takes a conical shape instead of a cylindrical one to


(a) (b)

(c)

Figure 2. The designed prosthetic artificial mechanical heart valve; (a) the designed conical caged ball valve; (b); the single hemispherical caged ball valve (c) the single-leaflet valve

increase control over the alignment of the central movement of the ball when the blood is flowing or when the valve is closed.

Figure 2b shows the designed single hemispherical caged ball artificial valve, Which works in the same way as the single-leaflet disc valve, but with better peripheral blood flow dynamics, ease, and rapid closing of the valve when the blood flow stops. Figure 2c shows the designed artificial single-leaflet prosthetic valve.

It consists of four perpendicular metallic or polymeric stents to hold a disc inward above the orifice for blood flow from the heart, and the opening angle of the disc concerning the valve ring ranges from 60 to 80 degrees. This type may show a complex dynamic flow, where the kinetic energy of the blood flow must first move the disc and then centralize the circumferential flow direction around it (8). These valves are similar to caged ball valves, except that the valve shutter is a disc (22-29). This valve allows a reduction in the annulus dimensions. The thickness of the disc valve ring is about 0.13D and thus its total length is reduced to 0.82D. It works rapidly and smoothly in the velocity field in the aorta (30).

The single-disc valve has been shown to operate at a much lower rate of clots than the bileaflet valve, and the open rate is greater (31). The structures of the pulmonary and aortic valves are similar, and both are made of three semilunar leaflets within the valve roots (5). Following the great success of interest treatments have grown in treatments to include the mitral valve. The first therapeutic mitral valve implanted in humans was approved by the Food and Drug Administration in 2013. Since then, several technologies have been developed and are currently being investigated (32). The structures of the pulmonary and aortic valves are similar, and both are made of three semilunar leaflets within the valve roots (5).

Figure 3 shows the designing steps of an artificial mitral tri-leaflet heart valve. This valve has been designed to work on the same principles as the previous valve, except that it allows the oxygen-rich blood to flow from the heart into the aorta. The prototype of this valve was produced from a short silicone rubber tube and a metal ring (33).

Then the design was improved by making it without a metal ring.Tricuspid valve disease is an unmet clinical need. Tricuspid valve disease is a common valvular with high patient and mortality rates. Yet tricuspid valve surgery remains the most dangerous of all valvular surgeries with an estimated inpatient mortality rate of 8.8%.

Due to the novelty of these techniques and the complex and variable anatomy of the tricuspid valve, procedural planning, and research based on 3D printing are of great value (32). Increased interest in tricuspid valve replacement has led to the advancement of 3D printing models of this valve (34-36). Tricuspid polymeric valves were developed that better resemble the original valve geometries and are more dynamically efficient. Figure 4 shows the designed tricuspid prosthetic aortic valve (TAV). It consists of three leaflets of equal shape and size (37).

This valve allows blood to flow from the aorta into the heart. It is a replacement for the bicuspid aortic valve which is a congenital defect since it is not able to completely stop the blood from leaking back into the heart. This can harden and not-open the valve, causing the heart to pump harder to bypass the aortic valve stenosis.

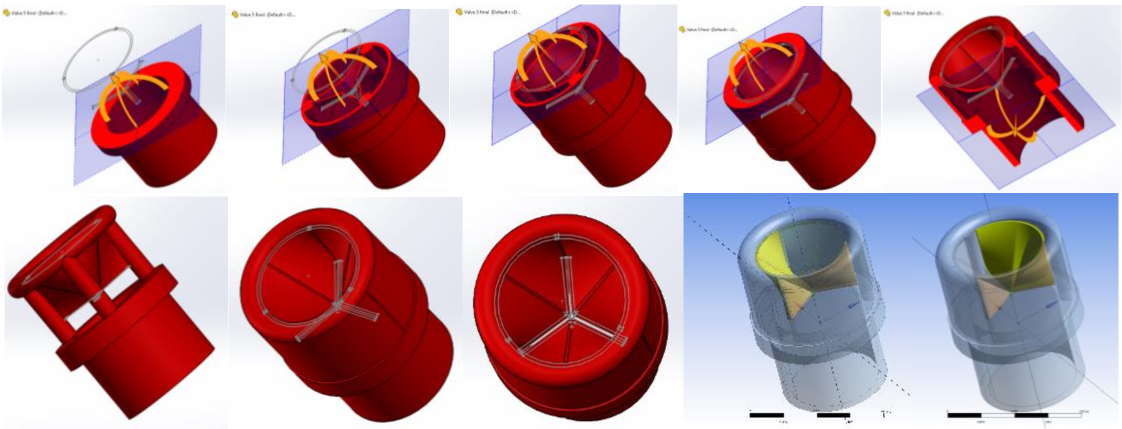


Figure 3. The designing steps mitral tri-leaflet heart artificial valve

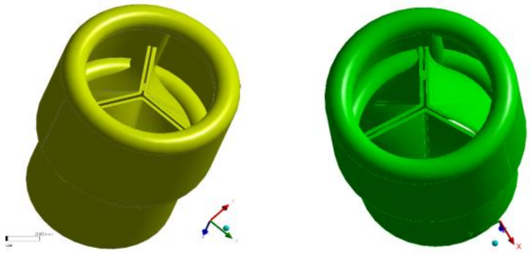


Figure 4. The designed tricuspid prosthetic aortic valve (TAV)

2. ٢. Mathematical Modeling of the Blood Flow

The blood in the human body contains liquid and solid parts. The liquid part of the blood is called plasma, generally made of water, protein, and salts, and over 50% of blood is plasma. For developing the Cardiovascular system equation, the artery material is assumed to be non-deformable, isotropic, tapered, without longitudinal movements, and incompressible with circular sections (38). To derive the equations governing the blood fluid flow, let us assume that the blood fluid flow is assumed to be axisymmetric and incompressible (39). However, under conditions of high blood flow, particularly in the ascending aorta, the laminar flow can be turbulent and disrupted and can be described as chaotic. Turbulence flow generates heat and is dissipated, as the increases the energy required to drive blood flow because turbulence increases the loss of energy as friction. Turbulence occurs when the velocity of flow becomes high enough that the flow lamina separates and when a critical Reynolds number (Re) is exceeded. Reynolds number is a way to predict, under ideal conditions, when turbulence will occur (40).

It's important to recognize that the Navier-Stokes equations are used based on the assumption of Newtonian fluid behavior, where the viscosity is constant regardless of the shear rate. However, in microcirculatory systems such as small branches and capillaries or regions with disturbed flow patterns like the blood flow through the mechanical artificial heart valves, blood exhibits non-Newtonian flow patterns as treated in the current work (39). The equations that describe the geometry of the artificial heart stenosis models in a bell shape can be expressed as in Equation 1:

$$\mathrm{R}(z) = -R_0\left(1 - ae^{-b(z-L)^2}\right) \tag{1}$$

where a and b are the dimensionless parameters; a = δ/R0 is the ratio of stenosed depth = 0.03–0.08 cm and the radius of a normal artery and b =(m/L0)2 is the square of the ratio of constant value m =1.4–1.8, and the half-length of the stenosis. Let the common parameters such as the radius of a normal artery (R0); R(z) is the radius of a stenosed artery; (L0) is the half-length of the stenosis; (L) is the half-length of an artery, and δ is the highest depth of the stenosis.

The blood flow model in an arterial segment is considered to be steady, incompressible, laminar, axisymmetric, and two-dimensional. The Boussinesq approximation is used to approximate the turbulent stresses existing by the turbulent dynamic viscosity (41). Then neglecting the orientation of gravity inside the body, the continuity and momentum equations for a non-Newtonian behavior, for the current blood flow study through the heart valve stenosis when the shear rate is less than 100 s-1. The equation of continuity can be written in the cylindrical coordinates system (r, θ, z) as follows (39, 42, 43):

$$\frac{\partial v_r}{\partial r} + \frac{v_r}{r} + \frac{\partial v_z}{\partial z} = 0. \tag{2}$$

The equation of momentum is:

$$v_r \frac{\partial v_z}{\partial r} + v_z \frac{\partial v_z}{\partial z} = -\frac{1}{p}\frac{\partial p}{\partial z} - \frac{1}{\rho}\left(\frac{1}{r}\frac{\partial (r\tau_{rz})}{\partial r} + \frac{\partial \tau_{zz}}{\partial z}\right) \tag{3}$$

$$v_r \frac{\partial v_r}{\partial r} + v_z \frac{\partial v_r}{\partial z} = -\frac{1}{p}\frac{\partial p}{\partial r} - \frac{1}{\rho}\left(\frac{1}{r}\frac{\partial (r\tau_{rr})}{\partial r} + \frac{\partial \tau_{rz}}{\partial z}\right) \tag{4}$$

where, τrz, τrr, and τzz are the shear stress of the model; vz is the axial velocity, p is the pressure of the fluid, vr is the radial velocity, and ρ is the density of the fluid, and:

$$\tau_{rz} = -k\gamma^{n-1}\left(\frac{\partial v_r}{\partial z} + \frac{\partial v_z}{\partial r}\right) \tag{5}$$

$$\tau_{zz} = -2k\gamma^{n-1}\left(\frac{\partial v_z}{\partial z}\right) \tag{6}$$

$$\tau_{rr} = -2k\gamma^{n-1}\left(\frac{\partial v_r}{\partial r}\right) \tag{7}$$

where k is the flow consistency index = 0.02 and:

$$\gamma = \sqrt{2\left[(\frac{\partial v_r}{\partial y})^2 + (\frac{v_r}{r})^2 + (\frac{\partial v_z}{\partial z})^2\right] + (\frac{\partial v_r}{\partial z} + \frac{\partial v_r}{\partial r})^2} \tag{8}$$

On the arterial wall, the axial velocity gradient is assumed to be zero, which means there is no shear rate of blood fluid along the axis and also there is no radial flow along the axis of the artery. Then the boundary conditions can be expressed as follow:

$$v_r(r,z) = 0, \quad \frac{\partial v_z}{\partial r}(r,z) = 0, \quad \tau_{rz} = 0, \quad on\, r = 0, \tag{9}$$

$$v_r(r,z) = 0, \quad v_z(r,z) = 0, \quad on\, r = R(z). \tag{10}$$

To immobilize the arterial wall, the radial coordinate transformation of governing equations can be used, where y=r/R(z). Then Equations 2 to 7 can be written as follow:

$$\frac{1}{R}\frac{\partial v_r}{\partial y} + \frac{v_r}{yR} + \frac{\partial v_z}{\partial z} - \frac{y}{R}\frac{dR}{dz}\frac{\partial v_z}{\partial y} = 0, \tag{11}$$

$$\left[-\frac{v_r}{R} + v_z\frac{y}{R}\frac{dR}{dz}\right]\frac{\partial v_z}{\partial y} - v_z\frac{\partial v_z}{\partial z} - \frac{1}{\rho}\frac{\partial p}{\partial z} - \frac{1}{\rho}\left[\frac{\tau_{yz}}{yR} + \frac{1}{R}\frac{\partial \tau_{yz}}{\partial z} - \frac{y}{R}\frac{dR}{dz}\frac{\partial \tau_{yz}}{\partial y}\right] = 0, \tag{12}$$

$$\tau_{yz} = -k\gamma^{n-1}\left(\frac{\partial v_r}{\partial z} - \frac{y}{R}\frac{dR}{dz}\frac{\partial v_r}{\partial y} + \frac{1}{R}\frac{\partial v_z}{\partial y}\right) \tag{13}$$

$$\tau_{zz} = -2k\gamma^{n-1}\left(\frac{\partial v_z}{\partial z} - \frac{y}{R}\frac{dR}{dz}\frac{\partial vz}{\partial y}\right), \tag{14}$$

where,

$$\gamma = \sqrt{2\left[(\frac{1}{R}\frac{\partial v_r}{\partial y})^2 + (\frac{v_r}{yR})^2 + (\frac{\partial v_z}{\partial z} - \frac{y}{R}\frac{dR}{dz}\frac{\partial v_z}{\partial y})^2\right] + (\frac{\partial v_r}{\partial z} - \frac{y}{R}\frac{dR}{dz}\frac{\partial v_r}{\partial y} + \frac{1}{R}\frac{\partial v_z}{\partial y})^2}. \tag{15}$$

The transformation initial and boundary conditions are (43):

$$v_r(y,z) = 0, \quad \frac{\partial v_z}{\partial y}(r,z) = 0, \quad \tau_{yz} = 0, \quad on\, y = 0, \tag{16}$$

$$v_r(y,z) = 0, \quad v_z(y,z) = 0, \quad on\, y = 1. \tag{17}$$

The radial velocity can be solved by multiplication of yR in Equation 11 and then integrated concerning y from 0 to 1, then:

$$\left[v_r\right]_0^1 = \left[y\frac{dR}{dz}v_z\right]_0^1 - \frac{R}{y}\int_0^1 y\frac{\partial v_z}{\partial z}dy - \frac{2}{y}\frac{dR}{dz}\int_0^1 yv_z dy. \tag{18}$$

By using the boundary condition (of Equation 16), Equation 18 will take the form:

$$\int_0^1 y\frac{\partial v_z}{\partial z}dy = -\frac{2}{R}\frac{dR}{dz}\int_0^1 yv_z dy. \tag{19}$$

Since R = R(z), Equation 19 can be re-written as:

$$\int_0^1 y\frac{\partial v_z}{\partial z}dy = -\int_0^1 y\frac{2}{R}\frac{dR}{dz}v_z dy. \tag{20}$$

By comparing both sides of the Equation 20, will give:

$$\frac{\partial v_z}{\partial z} = -\frac{2}{R}\frac{dR}{dz}v_z. \tag{21}$$

Substituting Equation 21 into Equation 11 and multiplying by yR, then:

$$y\frac{\partial v_r}{\partial y} + v_r = 2y\frac{dR}{dz}v_z + y^2\frac{dR}{dz}\frac{\partial v_z}{\partial y}. \tag{22}$$

Equation 22 can be re-written as follows:

$$\frac{\partial v}{\partial y}(yv_r) = \frac{\partial}{\partial y}\left(y^2\frac{dR}{dz}v_z\right). \tag{23}$$

Finally, the radial velocity is obtained with the form:

$$v_r = y\frac{dR}{dz}v_z. \tag{24}$$

To obtain the blood radial velocity (vr), it is required to plug vr in Equation 12, then equations of axial momentum can be reduced to:

$$-v_z\frac{\partial v_z}{\partial z} - \frac{1}{\rho}\frac{\partial p}{\partial z} - \frac{1}{\rho}\left[\frac{\tau_{yz}}{yR} + \frac{1}{R}\frac{\partial \tau_{yz}}{\partial y} + \frac{\partial \tau_{zz}}{\partial z} - \frac{y}{R}\frac{dR}{dz}\frac{\partial \tau_{yz}}{\partial y}\right] = 0. \tag{25}$$

The volumetric blood flow rate (Q) is given by the following equation:

$$Q = \int_0^R\int_0^{2\pi} v_z r\, d\theta dr = 2\pi\int_0^R v_r r\, dr. \tag{26}$$

And, using the radial coordinate transformation, r = yR, then:

$$Q = 2\pi R^2\int_0^1 v_r y\, dy. \tag{27}$$

Then, the radial velocity can be solved by discretization of Equation 24 as follows:

$$(v_r)_{i.j} = y_i\left(\frac{dR}{dz}\right)_j (v_z)_{i.j} \tag{28}$$

The volumetric flow rate can be approximated by discretizing Equation 27:

$$Q_j = 2\pi R_j^2 \Delta y \sum_{i=1}^{N+1} (v_z)_{i,j}\, y_i \quad for\ j = 1,2,...,M+1 \tag{29}$$

The resistance to the blood flow (λ) can be obtained as:

$$\lambda_j = \frac{\Delta p}{Q_j}. \tag{30}$$

To develop the mathematical model of blood pressure, Poiseuille's equation is considered to determine the relation between blood flow rate and pressure which is:

$$Q = \frac{\pi R^2}{8L v_z} P \tag{31}$$

Equation 25 is the mathematical model of the blood pressure in the human body. The required boundary condition and the values of the other parameters to solve these equations are; the pressure gradient, ∂ P /∂z = 40-100 mmHg, the Initial blood flow rate value of Q = 1 to 5.4 liter/minute, the Kinematic viscosity of blood, v = 0.035 cm^2/s and the density of blood ρ = 1.043 to 1.057 g/cm^3, the radius of a normal artery R0 = 0.1 cm, the constant dynamic viscosity (μ) of 3.71×10-3 Pa s, the blood density, initial value of S0 = 2.0 to 3.0 cm^3, and Q0 = 16.7 cm^3/s. For simplicity, the length of the artery model L = 14 cm, L0 = 3-8 cm, The mean Reynold number = 300, the artery systole duration for 0.3 seconds and diastole 0.5 seconds, and the period is 0.2 seconds (40, 44-46). Reynolds number and Re = of the blood flow is: Re = V * D * ρ/ μ .

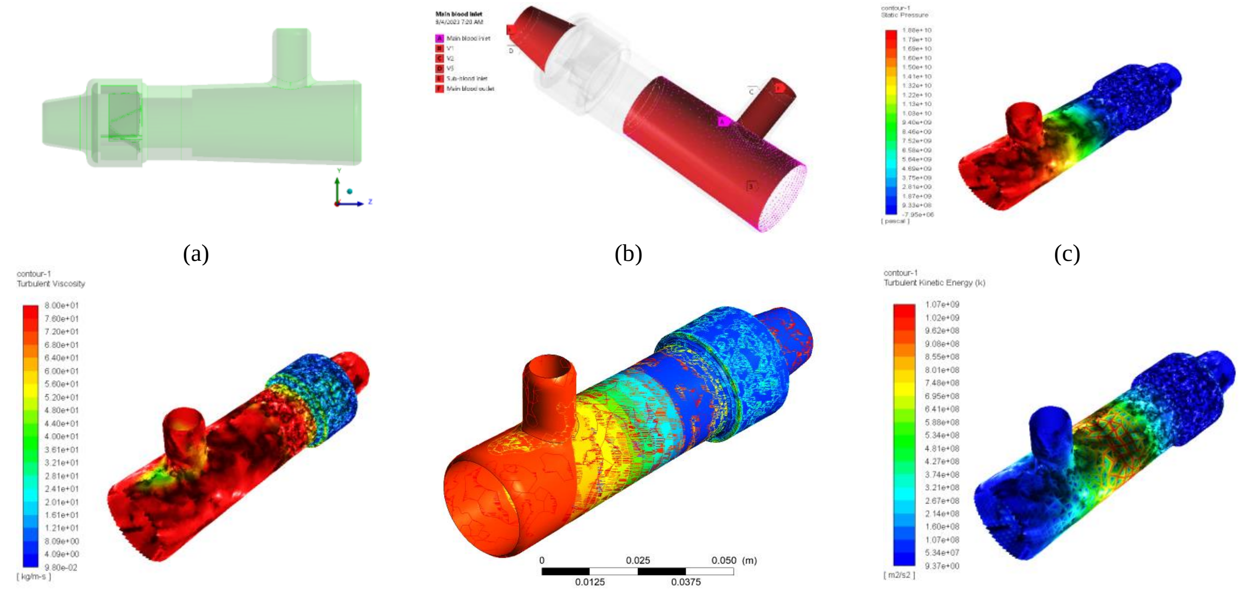
۳. RESULTS AND DISCUSSIONS

The equation of the mathematical model of blood flow rate can be solved using the Computational fluid dynamics (CFD) which is a branch of fluid mechanics that uses numerical analysis system, while the finite volume method (FVM) was used for discretization of flow governing equations.

Some general settings were set for the CFD analysis in the navigation pane to perform mesh-related activities including setting up the models for the CFD simulation's viscosity by activating a multiphase and viscose model; enabling the volume and viscosity of the fluid as laminar; and enabling the Energy Equation option from the Model list, the k-epsilon and K-equation turbulence model to expand the Viscous Model (45). In CFD analysis system, the Y+ value represents the distance from the first grid cell to the surface wall is an important parameter for determining the accuracy of the boundary layer thickness (46-48).

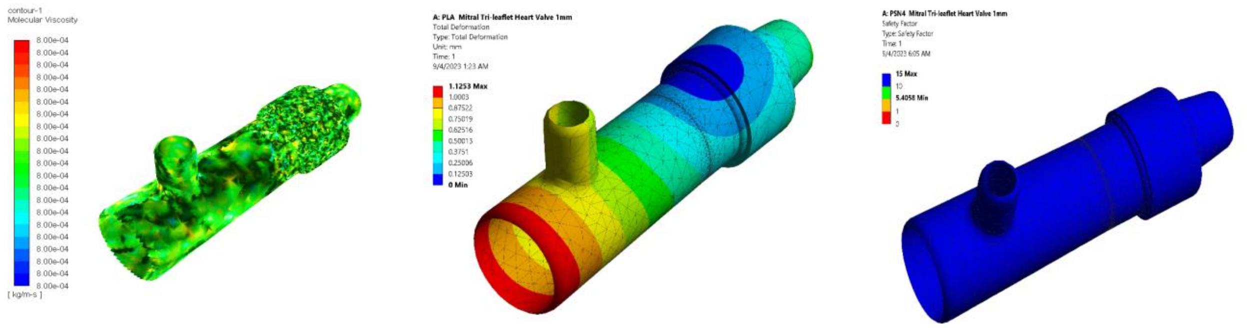
Figure 5 represents CFD solution and simulation of the of blood flow in the atria with a tricuspid mitral valve (1mm), using the high-strength and flexible PSN4 nanomaterial.

To determine the mechanical properties of the designed artificial valves, the response surface method (RSM), the finite element method (FEM) and the Design-Expert 13 were applied (49, 50). They were first modeled with the static structures analysis system. Then the meshing process was carried out with the additional refinement fragmentation for the moving parts of the valves, which are exposed to variable loads, as shown in Figure 6a.



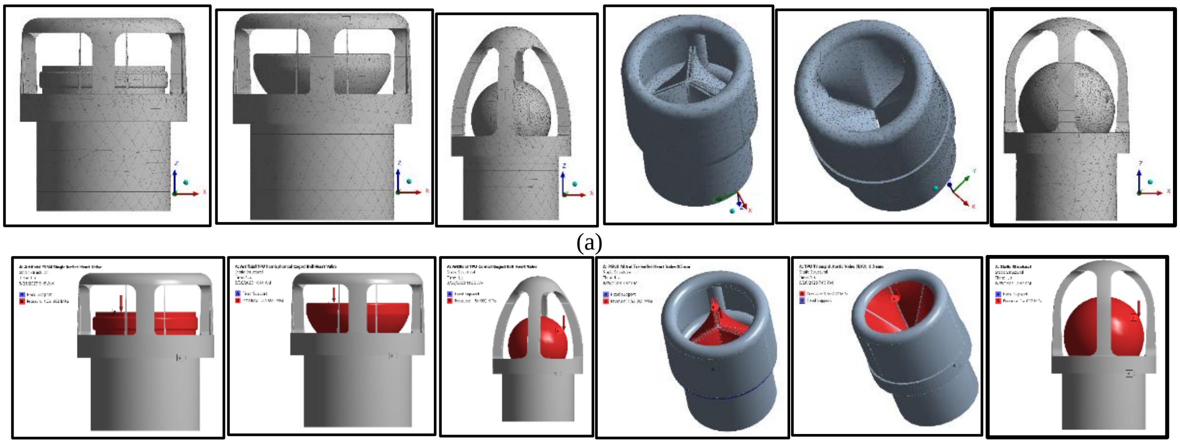
(a) (b) (c)

(d) (e) (f)



(g) (h) (i)

Figure 5. Simulation of blood flow in the atria; (a) The initial state of the artery; (b) Boundary Conditions; (c) Static pressure; (d);Turbulent viscosity (e) Eddy viscosity; (f); Turbulent kinetic energy (g) Molecular viscosity; (h) Total deformation; (i) Stress safety factor



(a)

(b)

Figure 6. Modeling of the designed artificial valves; (a) The meshing process; (b) The determined boundary conditions for the applied heart blood pressure from the hear

The structure mesh has been used for gridding of each designed artificial heart valve with the number of different cells and the mesh study performed according to the required number of cells were created for gridding (51).

Then the boundary conditions for each valve were determined, where a pressure equal to the blood flowing pressure from the heart, which is mmHg, was subjected on the surface of the valve moving parts, while the base area in the assembly completion with the heart was simply supported, as shown in the Figure 6b. Table 2 illustrates the mechanical properties of the designed artificial heart valves.

Figure 7 shows The strain energy simulations for selected types and the used materials of the designed artificial heart valves. Figures 8 to 10 show the simulations models of the strain energies, the total deformations, the equivalent stresses, the stress safety factor for selected types, and the used materials of the designed artificial heart valves.

To simulate of diastolic and systolic of the natural heart at each pulse cycle, the Goodman loading ratio criteria was used for the transient continuously and repeated applied blood pressure of +200 Pa for 0.5 second followed by an absorption pressure -50 Pa for the duration for 0.3 second (47). To analyze the results, the statistical Expert Systems 13.0 program was used.

Figure 11 shows the equivalent stress values resulting and calculated from applying blood pressure on the moving parts of these valves of each type of manufactured heart valve by using the Von Mises method.

From the figure, it is clear that the highest values of the equivalent stresses were found in valves with moving parts in a large range and three dimensions.

The highest values were produced in the mitral tri-leaflet valve (tri-leaflet Thick. = 0.5 mm) reached 14.13 MPa, followed by the tricuspid Aortic valve (TAV), (tricuspid thick. = 0.5 mm and 1 mm, respectively), while the equivalent stresses for the other types of the produced valves with the simple surface movement were less than 2 MPa, and the lowest value of stress was produced in the single-leaflet type valve, which reached 0.87 MPa.

TABLE 2. The fatigue properties of the designed artificial heart valves

Artificial Heart Valve Type	Filament Material	Strain Energy (10-5 mJ)	Total Deformation (mm)	Equivalent stress (MPa)	Stress safety factor (Min.)	Artificial Heart Valve Type	Filament Material	Strain Energy (10-5 mJ)	Total Deformation (mm)	Equivalent stress (MPa)	Stress safety factor (Min.)
Caged Ball	PETG	1.866	0.001	1.499	1.601	**Tricuspid Aortic Valve (TAV), (Tricuspid Thick. = 0.5 mm)**	PETG	178.100	0.585	11.077	0.217
	PLA	1.460	0.001		1.801		PLA	139.250	0.457		0.244
	Nylon	20.553	0.015		2.001		Nylon	1960.900	6.441		0.271
	TPE	3.347	0.025		4.700		TPE	382.370	1.256		0.632
	TPU	4.008	0.032		5.337		TPU	4858.600	9.959		0.722
	SIBSTAR 103T	0.677	0.001		5.671		SIBSTAR 103T	64.589	0.212		0.767
	PSN4	0.300	0.0002		12.00		PSN4	28.642	0.094		1.625
Conical Caged Ball	PETG	0.834	0.0010	1.618	1.484	**Tricuspid Aortic Valve (TAV), (Tricuspid Thick. =1.0 mm)**	PETG	11.281	20.841	13.028	0.184
	PLA	0.649	0.0008		1.669		PLA	8.824	0.124		0.207
	Nylon	9.192	0.0113		1.855		Nylon	124.26	1.745		0.230
	TPE	1.792	0.0022		4.328		TPE	24.231	0.340		0.480
	TPU	1.792	0.0280		4.946		TPU	307.890	43.247		0.614
	SIBSTAR 103T	0.303	0.0004		5.26		SIBSTAR 103T	4.093	0.058		0.652
	PSN4	0.134	0.0002		11.12		PSN4	1.815	0.026		1.382
Single-leaflet	PETG	20.405	0.0003	0.150	15	**Mitral Tri-leaflet Valve (Tri-leaflet Thick. = 0.5 mm)**	PETG	84.590	0.629	14.132	0.170
	PLA	18.816	0.0003		15		PLA	66.168	0.492		0.191
	Nylon	2.650	0.0037		15		Nylon	931.790	6.926		0.212
	TPE	51.650	0.0007		15		TPE	181.700	1.351		0.495
	TPU	6.565	0.0090		15		TPU	2308.700	11.160		0.566
	SIBSTAR 103T	872.780	0.0001		15		SIBSTAR 103T	30.692	0.228		0.602
	PSN4	387.080	0.00005		15		PSN4	13.610	0.101		1.274
Single Hemispherical Leaflet	PETG	0.845	0.0010	0.870	2.759	**Mitral Tri-leaflet Valve (Tri-leaflet Thick. = 1.0 mm)**	PETG	62.986	0.108	7.339	0.327
	PLA	0.661	0.0008		3.103		PLA	49.269	0.085		0.368
	Nylon	9.312	0.0113		3.448		Nylon	693.810	1.194		0.409
	TPE	1.816	0.0022		8.046		TPE	135.290	0.233		0.954
	TPU	23.044	0.0280		9.195		TPU	1719.100	2.959		1.090
	SIBSTAR 103T	0.308	0.0004		9.77		SIBSTAR 103T	22.854	0.039		1.158
	PSN4	0.136	0.0002		15		PSN4	10.134	0.0170		2.453

Figure 12 shows the relationship between the produced artificial heart valve type and the filament material with the strain energy expended when blood pressures of 120 mmHg are applied on the moving parts of the manufactured heart valves from the heart to the arteries or from the veins to the heart to open or close the blood flow during the process of diastole and contraction of the heart.

This figure shows that the energy expended is directly proportional to the strength and flexibility of the used material, and the energy expenditure rates decrease when using highly flexible materials such as TPE and PSN4. The values of this energy also increase when the area of the moving parts of the valve increases, especially when the closure process is facing the blood flow, as is the case with the use of a tricuspid Aortic valve (TAV), (tricuspid thick. = 0.5 mm and 1 mm, respectively). The amount of energy increased when using the mitral tri-leaflet valve (tri-leaflet Thick. = 0.5 mm and 1 mm, respectively). The least energy was spent when using simple movement valves in one direction, which are the conical caged ball, the single hemispherical leaflet, the caged ball, and then the single leaflet valves, respectively. In general, the highest calculated strain energy does not exceed 0.049 mJ.

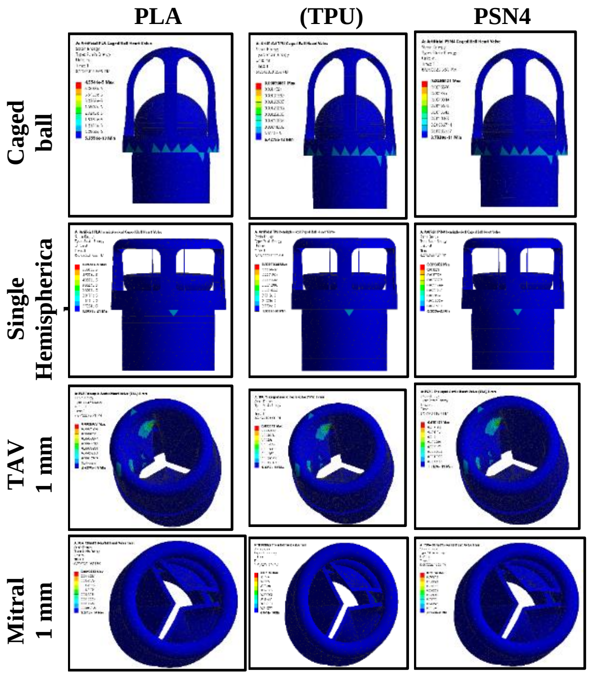


Figure 7. The strain energy simulations for selected types and the used materials of the designed artificial heart valves

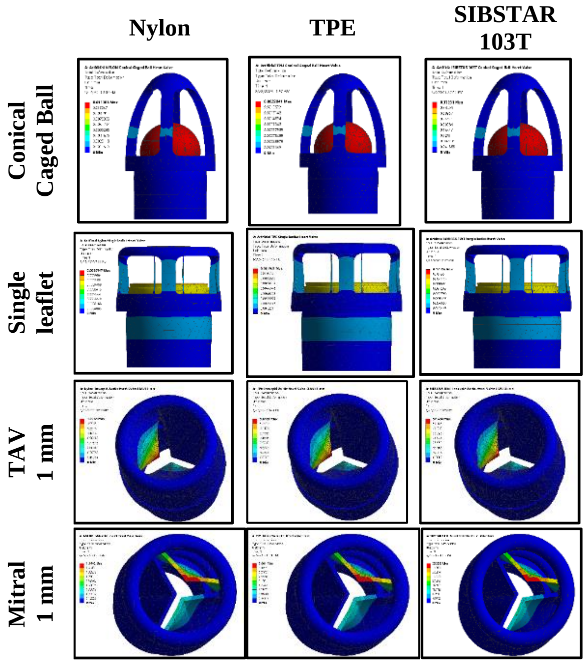


Figure 8. The total deformation simulations selected types and the used materials of the designed artificial heart valves

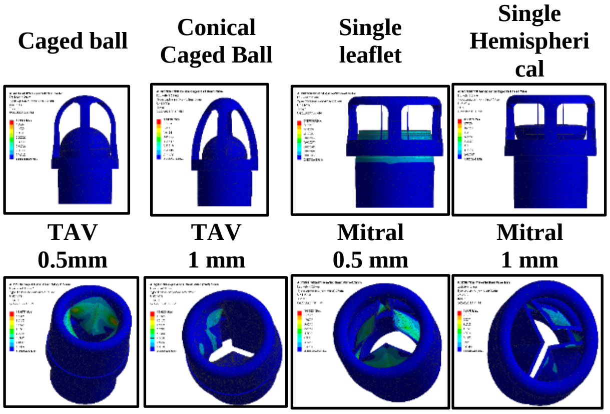


Figure 9. The equivalent stress for all the types and the used materials of the designed artificial heart valves

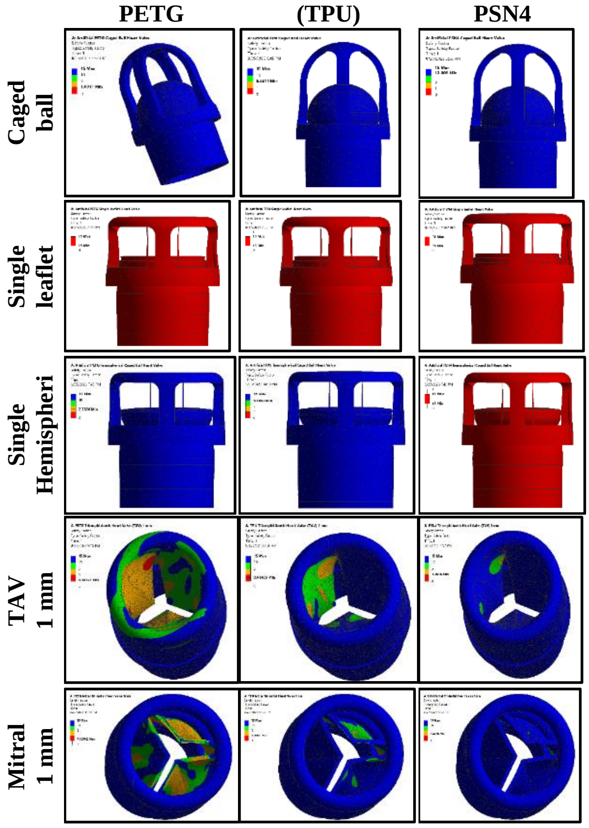


Figure 10. The stress safety factor simulations for selected types and the used materials of the designed artificial heart valves

Figure 13 shows the relationship between the produced heart valve type and the used material with the total deformation produced in the valve structure under the influence of applying the same blood pressure above. The figure shows that the highest total deformation in (mm), occurs when using thermoplastic polyurethane (TPU), followed by thermoplastic polyurethane

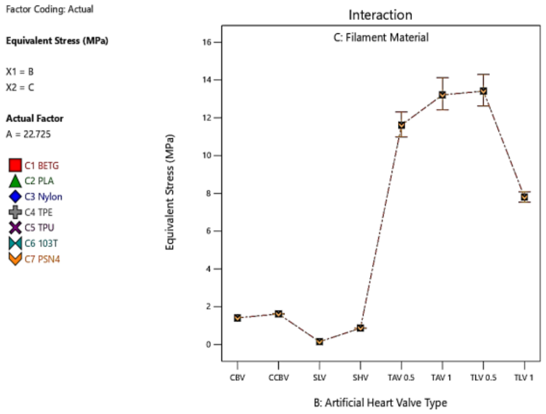


Figure 11. The equivalent stress for each produced artificial heart valve type

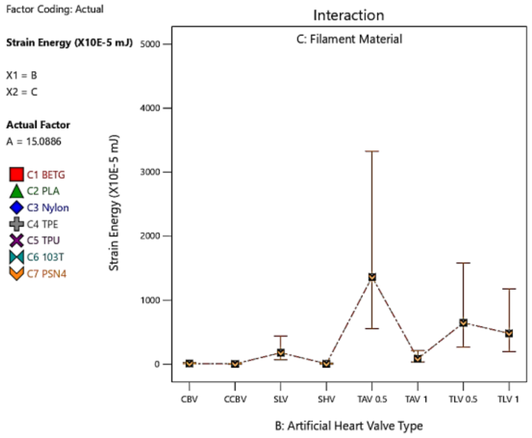


Figure 12. The expended strain energy for each designed artificial heart valve types

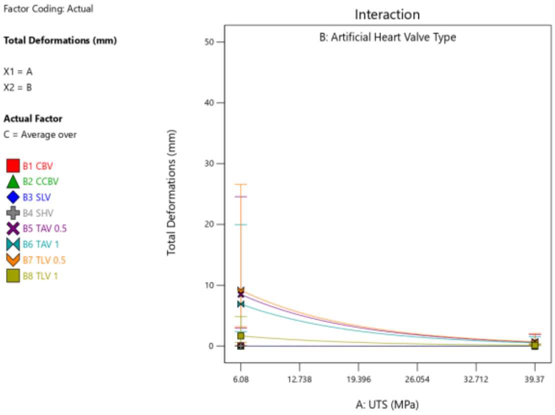


Figure 13. The total deformation produced in each artificial heart valve structure under the influence of applying the blood pressure

elastomer (TPE), nylon, polyethylene terephthalate glycol (PETG) and polyester PLA (Polylactic acid). Also, the lowest rates of deformation under the influence of blood pressure were observed when using materials of superior flexibility and strength, such as the used nanocomposite polyetherimide/silicone rubber with nano silica (PSN4), reaching from 5x10-5 to 0.1 mm, followed by the polystyrene- isobutylene-b-styrene (SIBSTAR103T) nanostructured rubber linear triblock.

Figure 14 shows the relationship between the material and type of the produced artificial heart valve with the stress safety factor, where the obtained values of this factor decrease with the complexity of the movement of the moving parts of the valve.

The highest rates were recorded when using the mitral tri-leaflet valve (tri-leaflet thick. = 1.0 mm) and reached 2.45 when using the high-strength and flexible PSN4 nanocomposite material. This value decreases to 1.27 when using the same material and type of valve but with a tri-leaflet thick. = 0.5 mm.

It is noted here that the PSN4 material is the only suitable for the manufacture of this artificial valve type, as well as the tricuspid aortic valve (TAV), where the other six tested materials failed because the stress safety factors are less than 1.

For the other types of manufactured valves, it was noted that the highest safety factor reached 15 was obtained when using the single-leaflet type valve, where all materials used in this work are suitable for the production of this type of valve because its movement is simple, flat, and in one direction. The principle of action of the tri-leaflet mechanical heart valve (BMHV) is that it has three semicircular leaflets prisms connected by small hinges with a solid, semi-circular valve ring instead of a single disc, and patients using this valve require anticoagulants for life (18). It has a much lower opening rate than the single-disc leaflet (33). It has a long service life and reliable performance and is of great importance in the treatment of valvular heart patients, but it still has defects such as thrombocytopenia, hemolysis, and thromboembolism (16, 20). The thickness of the leaflets is about 0.04D, and each leaflet can be opened up to 85

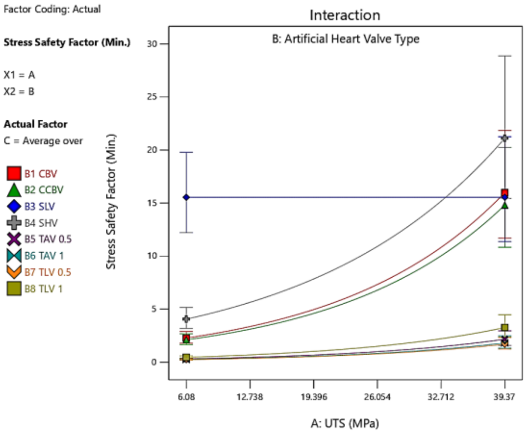


Figure 14. The stress safety factor for all the designed artificial heart valve types

degrees to obtain a rapid distribution of the flow velocity field of blood in the direction of the valve with less thrombosis (13, 21, 52).

The highest safety factor rates were observed after that when using the single hemispherical leaflet type valve, then the conical caged ball and the caged ball type, respectively. It can be concluded that the best materials for manufacturing these four types of valves are the PSN4, followed by the SIBSTAR103T, TPU, and TPE.

It is not recommended to use PETG, PLA, and nylon materials in the manufacture of any artificial heart valves, due to their lack of strength, flexibility, and high brittleness, especially PETG and PLA materials. SIBS material reinforced with woven polyester fabric or carbon nanotubes is also a promising material for the production of artificial heart valves. It has good mechanical properties, high thermal properties, and oxidative stability and is stable, biocompatible, and hematologically compatible making it an ideal candidate for medical applications (53, 54).

4. CONCLUSIONS

- It can be concluded that the best materials for manufacturing these four types of valves are the PSN4, followed by SIBSTAR103T, TPU, and TPE. The use of PETG, PLA, and nylon materials is not recommended in the manufacture of any prosthetic heart valves, due to their lack of strength, flexibility, and high brittleness, especially for PETG and PLA materials.
- It was also noted here that PSN4 is the only material suitable for the manufacture of mitral tri-leaflet and tricuspid mitral valve artificial valves. For other types of valves manufactured with a single leaflet, high safety stress factors were obtained because their movement is simple, flat, and in one direction, where the highest values were observed when testing a single hemispherical leaflet type valve, then the conical caged ball and the caged ball type, respectively .
- The highest values of the equivalent stress due to the applied blood pressure on the moving parts on each type of manufactured heart valve occur in valves with three dimensions moving parts, reached in the mitral tri-leaflet valve reached 14.13 MPa, followed by the tricuspid aortic valve. The equivalent stresses for other types of valves produced with simple surface action were lower than 2 MPa.
- The strain energy that is expended during the process of diastole and systole was found to be directly proportional to the strength and flexibility of the materials used. The energy consumption rates decrease when using highly elastic materials such as TPE and PSN4. The values of this energy also increase with an increase in the area of the moving parts of the valve, especially when faced with the process of closing blood flow, as with the use of the tricuspid aortic valve (TAV). The least values of energies were expended when using simple movement valves in one direction, including the single hemispherical leaflet, the caged ball, and then the single leaflet valves, respectively.
- The highest total deformation resulted in the valve body when using TPU, followed by TPE, nylon, PETG, and PLA, while the lowest deformation rates were observed when using PSN4, which ranged from 5x105 to 0.1 mm, followed by SIBSTAR103 nanostructured rubber.
- The obtained values of stress safety factors were decreased with the complexity of the movement for the moving parts of the valve. The highest rates were recorded when using the tricuspid mitral valve, reaching 2.45 when using the high-strength and flexible PSN4 nanomaterial .
- Additional future work is required to do mathematical modeling with numerical simulation of pressure distribution, blood flow velocity, and variable viscosity for all types of artificial heart valves presented in this paper.
- Working on the production of these types of artificial heart valves using the 3D bioprinting method, based on the successful application of tissue engineering in the development of heart valves to produce valves with mechanical properties similar to the original valves.

5. REFERENCES

1. Mitra M. Editorial on Advances in Artificial Heart. Ann Heart. 2018;3(1):51-2. 10.36959/652/390
2. Yoganathan AP, Fogel M, Gamble S, Morton M, Schmidt P, Secunda J, et al. A new paradigm for obtaining marketing approval for pediatric-sized prosthetic heart valves. The Journal of thoracic and cardiovascular surgery. 2013;146(4):879-86. 10.1016/j.jtcvs.2013.04.016
3. Ebad M, Vahidi B. In Silico Analysis of Stem Cells Mechanical Stimulations for Mechnoregulation Toward Cardiomyocytes. International Journal of Engineering, Transactions B: Applications, . 2022;35(11):2229-37. 10.5829/IJE.2022.35.11B.18
4. Hasan A, Soliman S, El Hajj F, Tseng Y-T, Yalcin HC, Marei HE. Fabrication and in vitro characterization of a tissue engineered PCL-PLLA heart valve. Scientific reports. 2018;8(1):8187. 10.1038/s41598-018-26452-y
5. Oveissi F, Naficy S, Lee A, Winlaw D, Dehghani F. Materials and manufacturing perspectives in engineering heart valves: a review. Materials Today Bio. 2020;5:100038. https://doi.org/10.1016/j.mtbio.2019.100038
6. Cook JA, Shah KB, Quader MA, Cooke RH, Kasirajan V, Rao KK, et al. The total artificial heart. Journal of thoracic disease. 2015;7(12):2172. 10.3978/j.issn.2072-1439.2015.10.70
7. Otto CM, Nishimura RA, Bonow RO, Carabello BA, Erwin III JP, Gentile F, et al. 2020 ACC/AHA guideline for the

management of patients with valvular heart disease: executive summary: a report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines. Journal of the American College of Cardiology. 2021;77(4):450-500. 10.1161/CIR.0000000000000932

8. Pibarot P, Dumesnil JG. Prosthetic heart valves: selection of the optimal prosthesis and long-term management. Circulation. 2009;119(7):1034-48. 10.1161/CIRCULATIONAHA.108.778886
9. Zhang DY, Lozier J, Chang R, Sachdev V, Chen MY, Audibert JL, et al. Case study and review: treatment of tricuspid prosthetic valve thrombosis. International journal of cardiology. 2012;162(1):14-9. 10.1016/j.ijcard.2011.09.081
10. Durko AP, Head SJ, Pibarot P, Atluri P, Bapat V, Cameron DE, et al. Characteristics of surgical prosthetic heart valves and problems around labeling: a document from the European Association for Cardio-Thoracic Surgery (EACTS)—the Society of Thoracic Surgeons (STS)—American Association for Thoracic Surgery (AATS) valve labelling task force. The Journal of thoracic and cardiovascular surgery. 2019;158(4):1041-54. 10.1093/ejcts/ezz034
11. Hasan A, Ragaert K, Swieszkowski W, Selimović Š, Paul A, Camci-Unal G, et al. Biomechanical properties of native and tissue engineered heart valve constructs. Journal of biomechanics. 2014;47(9):1949-63. https://doi.org/10.1016/j.jbiomech.2013.09.023
12. Ciolacu DE, Nicu R, Ciolacu F. Natural polymers in heart valve tissue engineering: Strategies, advances and challenges. Biomedicines. 2022;10(5):1095. https://doi.org/10.3390/biomedicines10051095
13. Rajashekar P. Development of mechanical heart valves-an inspiring tale. Journal of the Practice of Cardiovascular Sciences. 2015;1(3):289-93. 10.4103/2395-5414.177309
14. Rezvova MA, Klyshnikov KY, Gritskevich AA, Ovcharenko EA. Polymeric heart valves will displace mechanical and tissue heart valves: a new era for the medical devices. International Journal of Molecular Sciences. 2023;24(4):3963. https://doi.org/10.3390/ijms24043963
15. Gholami P, Kouchakzadeh MA, Farsi M. A continuum damage mechanics-based piecewise fatigue damage model for fatigue life prediction of fiber-reinforced laminated composites. International Journal of Engineering, Transactions C: Aspects. 2021;34(6):1512-22. 10.5829/IJE.2021.34.06C.15
16. Li W-q, Gao Z-x, Jin Z-j, Qian J-y. Transient study of flow and cavitation inside a bileaflet mechanical heart valve. Applied Sciences. 2020;10(7):2548. https://doi.org/10.3390/app10072548
17. Ghanbari J, Dehparvar A, Zakeri A. Design and analysis of prosthetic heart valves and assessing the effects of leaflet design on the mechanical attributes of the valves. Frontiers in Mechanical Engineering. 2022;8:764034. https://doi.org/10.3389/fmech.2022.764034
18. Nair K, Muraleedharan C, Bhuvaneshwar G. Developments in mechanical heart valve prosthesis. Sadhana. 2003;28:575-87. 10.1007/BF02706448
19. Dumont K. Experimental and numerical modeling of heart valve dynamics: Ghent University; 2004.
20. Dasi LP, Simon HA, Sucosky P, Yoganathan AP. Fluid mechanics of artificial heart valves. Clinical and experimental pharmacology and physiology. 2009;36(2):225-37. 10.1111/j.1440-1681.2008.05099.x
21. Jóźwik K, Witkowski D. Artificial heart valve designs. Mechanics and Mechanical Engineering. 2000;4(1):63--70.
22. Avinc O, Khoddami A. Overview of poly (lactic acid)(PLA) fibre: Part I: production, properties, performance, environmental impact, and end-use applications of poly (lactic acid) fibres. Fibre Chemistry. 2009;41(6):391-401. 10.1007/s10692-010-9213-z
23. Subbarao CV, Reddy YS, Inturi V, Reddy MI, editors. Dynamic mechanical analysis of 3D printed PETG material. IOP Conference Series: Materials Science and Engineering; 2021: IOP Publishing. 10.1088/1757-899X/1057/1/012031
24. Wang J, Yang B, Lin X, Gao L, Liu T, Lu Y, et al. Research of TPU materials for 3D printing aiming at non-pneumatic tires by FDM method. Polymers. 2020;12(11):2492. https://doi.org/10.3390/polym12112492
25. Storey RF, Baugh D, Choate K. Poly (styrene-b-isobutylene-b-styrene) block copolymers produced by living cationic polymerization: I. Compositional analysis. Polymer. 1999;40(11):3083-90. https://doi.org/10.1016/S0032-3861(98)00520-5
26. Pinchuk L, Wilson GJ, Barry JJ, Schoephoerster RT, Parel J-M, Kennedy JP. Medical applications of poly (styrene-block-isobutylene-block-styrene)(“SIBS”). Biomaterials. 2008;29(4):448-60. https://doi.org/10.1016/j.biomaterials.2007.09.041
27. Mishra R, Rai J. Polyetherimide (PEI)/silicone rubber composite reinforced with nanosilica particles. International Journal of Scientific and Technology Research. 2016;5(3):176-80.
28. Nikoohemmat MA, Mazaheri H, Joshaghani A, Joudaki E. Investigation on Physical and Mechanical Properties of High Density Polyethylene (PE100) Using Novel Catalyst. International Journal of Engineering, Transactions B: Applications,. 2022;35(11):2205-12. 10.5829/IJE.2022.35.11B.15
29. Hirai S, Fukunaga S, Sueshiro M, Watari M, Sueda T, Matsuura Y. Assessment of a new silicone tri-leaflet valve seamlessly assembled with blood chamber for a low-cost ventricular assist device. Hiroshima journal of medical sciences. 1998;47(2):47-55.
30. Szykiedans K, Credo W, Osiński D. Selected mechanical properties of PETG 3-D prints. Procedia Engineering. 2017;177:455-61. https://doi.org/10.1016/j.proeng.2017.02.245
31. Wu Y, Gregorio R, Renzulli A, Onorati F, De Feo M, Grunkemeier G, et al. Mechanical heart valves: are two leaflets better than one? The Journal of thoracic and cardiovascular surgery. 2004;127(4):1171-9. 10.1016/j.jtcvs.2003.08.030
32. Harb SC, Rodriguez LL, Vukicevic M, Kapadia SR, Little SH. Three-dimensional printing applications in percutaneous structural heart interventions. Circulation: Cardiovascular Imaging. 2019;12(10):e009014. 10.1161/CIRCIMAGING.119.009014
33. Jin Y, Rao A, Brinkman W, Choi T-Y. 3D printing-assisted energy loss testing of artificial aortic heart valves. arXiv preprint arXiv:191011191. 2019. https://doi.org/10.48550/arXiv.1910.11191
34. Vukicevic M, Filippini S, Little SH. Patient-specific modeling for structural heart intervention: role of 3D printing today and tomorrow CME. Methodist DeBakey cardiovascular journal. 2020;16(2):130. 10.14797/mdcj-16-2-130
35. Altaani HA, Jaber S. Tricuspid valve replacement, mechnical vs. biological valve, which is better? International cardiovascular research journal. 2013;7(2):71.
36. C. S. Samuel and K. S. Candice. Bicuspid Aortic Valve Disease London and Toronto, Ontario, Canada. Journal of the American College of Cardiology,. 2010;55(25):2789–800. 10.1016/j.jacc.2009.12.068
37. Yousefi M, Safikhani H, Jabbari E, Tahmsbi V. Numerical modeling and optimization of respirational emergency drug delivery device using computational fluid dynamics and response surface method. International Journal of Engineering, Transactions B: Applications 2021;34(2):547-55. 10.5829/IJE.2021.34.02B.28

38. Karthik A, Kumar PP, Radhika T. A mathematical model for blood flow accounting for the hematological disorders. Computational and Mathematical Biophysics. 2022;10(1):184-98. https://doi.org/10.1515/cmb-2022-0136

39. Kabir MA, Sultana K, Uddin MA. Performance of k-ω and k-ε Model for Blood Flow Simulation in Stenosed Artery. Ganit: Journal of Bangladesh Mathematical Society. 2020;40(2). https://doi.org/10.3329/ganit.v40i2.51314

40. Kabir MA, Alam MF, Uddin MA. A numerical study on the effects of Reynolds number on blood flow with spiral velocity through regular arterial stenosis. Chiang Mai J Sci. 2018;45(6):2515-27.

41. Nanda S, Mallik BB, Majumder SD, Karthick RK, Suman S, Sonkar S. Mathematical Modelling of Pulsatile Flow of Non-Newtonian Fluid Through a Constricted Artery. Mathematical Modelling of Engineering Problems. 2021;8(3). https://doi.org/10.18280/mmep.080320

42. Owasit P, Sriyab S. Mathematical modeling of non-Newtonian fluid in arterial blood flow through various stenoses. Advances in Difference Equations. 2021;2021(1):1-20. https://doi.org/10.1186/s13662-021-03492-9

43. Nikolić A, Topalović M, Simić V, Blagojević M. BLOOD FLOW IN ARTERIAL BIFURCATION CALCULATED BY TURBULENT FINITE ELEMENT MODEL. Journal of the Serbian Society for Computational Mechanics/Vol. 2021;15(2):79-92. 10.24874/jsscm.2021.15.02.08

44. Labadin J, Ahmadi A, editors. Mathematical modeling of the arterial blood flow. Proceedings of the 2nd IMT-GT Regional Conference on Mathematics, Statistics and Applications, Universiti Sains Malaysia, Penang; 2006.

45. Rahman MS, Haque MA, editors. Mathematical modeling of blood flow. 2012 International Conference on Informatics, Electronics & Vision (ICIEV); 2012: IEEE. 10.1109/ICIEV.2012.6317446

46. Chakravarty S, Mandal P. Mathematical modelling of blood flow through an overlapping arterial stenosis. Mathematical and computer modelling. 1994;19(1):59-70.

47. Ali SM, Ali ZJ, Abd MM, editors. Design and Modeling of a Soft Artificial Heart by Using the SolidWorks and ANSYS. IOP Conference Series: Materials Science and Engineering; 2020: IOP Publishing. 10.1088/1757-899X/671/1/012062

48. Jalili P, Kazerani K, Jalili B, Ganji D. Investigation of thermal analysis and pressure drop in non-continuous helical baffle with different helix angles and hybrid nano-particles. Case Studies in Thermal Engineering. 2022;36:102209. https://doi.org/10.1016/j.csite.2022.102209

49. Rajaee M, Hosseinipour S, Jamshidi Aval H. Multi-objective optimization of HMGF process parameters for manufacturing AA6063 stepped tubes using FEM-RSM. International Journal of Engineering, Transactions B: Applications 2021;34(5):1305-12. 10.5829/IJE.2021.34.05B.25

50. Jalili P, Ganji D, Nourazar S. Investigation of convective-conductive heat transfer in geothermal system. Results in Physics. 2018;10:568-87. https://doi.org/10.1016/j.rinp.2018.06.047

51. Daebritz SH, Sachweh JrS, Hermanns B, Fausten B, Franke A, Groetzner J, et al. Introduction of a flexible polymeric heart valve prosthesis with special design for mitral position. Circulation. 2003;108(10_suppl_1):II-134-II-9. 10.1161/01.cir.0000087655.41288.dc

52. Khalili F, Gamage P, Mansy HA. Hemodynamics of a bileaflet mechanical heart valve with different levels of dysfunction. arXiv preprint arXiv:171111153. 2017. 10.15406/jabb.2017.02.00044

53. Makarevich MI, Nikishau PA, Berezianko IA, Glushkova TV, Rezvova MA, Ovcharenko EA, et al. Aspects of the synthesis of poly (styrene-block-isobutylene-block-styrene) by TiCl4-co-initiated cationic polymerization in open conditions. Macromol. 2021;1(4):243-55. https://doi.org/10.3390/macromol1040017

54. Farah S, Anderson DG, Langer R. Physical and mechanical properties of PLA, and their functions in widespread applications—A comprehensive review. Advanced drug delivery reviews. 2016;107:367-92. 10.1016/j.addr.2016.06.012

Persian Abstract

چکیده

جایگزینی دریچه قلب یک بار سلامتی بزرگ است و میلیون‌ها نفر در سراسر جهان به آن نیاز دارند، که نیاز مستمر به کشف و ساخت جایگزین‌های مصنوعی مؤثرتر و دائمی‌تر را می‌طلبد. در کار حاضر، مدل های منحصر به فرد هشت دریچه قلب مصنوعی با استفاده از هفت ماده مصنوعی و نانوکامپوزیت طراحی و مورد بررسی قرار گرفت. شیرهای طراحی شده برای تعیین بهترین طرح ها و مواد از نظر دوام، انعطاف پذیری، مصرف انرژی و عملکرد مورد بررسی قرار گرفتند. طراحی و شبیه سازی دریچه های مصنوعی انجام شد و برای اعتبارسنجی و بهبود عملکرد بیومکانیکی با استفاده از روش سطح پاسخ (RSM) و سیستم خبره طراحی ۱۳. بالاترین مقادیر تنش معادل به دلیل فشار خون اعمال شده بر روی قطعات متحرک در هر نوع دریچه قلب ساخته شده در دریچه هایی با قطعات متحرک سه بعدی که در دریچه سه لنگه میترال ۱٤.۱۳ مگاپاسکال و به دنبال آن دریچه آئورت سه لتی به دست می آید، رخ می دهد. تنش های معادل برای انواع دیگر شیرهای تولید شده با عملکرد سطحی ساده کمتر از ۲ مگاپاسکال بود. انرژی کرنشی که در طی فرآیند دیاستول و سیستول صرف می‌شود مشخص شد که با قدرت و انعطاف‌پذیری مواد مورد استفاده رابطه مستقیم دارد. نرخ مصرف انرژی هنگام استفاده از مواد بسیار الاستیک مانند TPE و PSN4 کاهش می یابد. مقادیر این انرژی نیز با افزایش سطح قسمت‌های متحرک دریچه افزایش می‌یابد، به‌ویژه هنگامی که با فرآیند بسته شدن جریان خون مواجه می‌شوید، مانند استفاده از دریچه آئورت سه لتی (TAV). بیشترین تغییر شکل کل در بدنه دریچه هنگام استفاده از TPU و به دنبال آن TPE، نایلون، PETG و PLA، در حالی که کمترین نرخ تغییر شکل در هنگام استفاده از PSN4 مشاهده شد که از x105۵ تا ۰.۱ میلی‌متر و به دنبال آن لاستیک نانوساختار SIBSTAR103 قرار داشت. مقادیر به‌دست‌آمده عوامل ایمنی تنش با پیچیدگی حرکت برای قسمت‌های متحرک شیر کاهش می‌یابد. بالاترین میزان در هنگام استفاده از دریچه میترال تریکوسپید ثبت شد و در هنگام استفاده از نانومواد PSN4 با مقاومت بالا و انعطاف پذیر به ۲.٤٥ رسید. می توان نتیجه گرفت که بهترین مواد برای ساخت این چهار نوع شیر، PSN4 و به دنبال آن SIBSTAR103T، TPU و TPE هستند. استفاده از مواد PETG، PLA و نایلون در ساخت هیچ گونه دریچه مصنوعی قلب به دلیل عدم استحکام، انعطاف پذیری و شکنندگی بالا به خصوص برای مواد PETG و PLA توصیه نمی شود. همچنین در اینجا ذکر شد که PSN4 تنها ماده مناسب برای ساخت دریچه های مصنوعی دریچه سه لفت و سه لتی میترال است. برای انواع دیگر دریچه‌های تولید شده با تک لت، فاکتورهای استرس ایمنی بالا به دست آمد زیرا حرکت آنها ساده، مسطح و در یک جهت است، جایی که بالاترین مقادیر را هنگام آزمایش یک دریچه از نوع منفرد نیم‌کره مشاهده شد، سپس توپ مخروطی در قفس مشاهده شد. و به ترتیب نوع توپ در قفس.

IJE TRANSACTIONS B: Applications Vol. 37 No. 02, (February 2024) 252-267

International Journal of Engineering

Journal Homepage: www.ije.ir

Design and Analysis of a Power Quality Improvement System for Photovoltaic Generation Based on LCL-Type Grid Connected Inverter

M. Hosseinpour*[a], A. Dastgiri[a], M. Shahparasti[b]

[a] *Department of Electrical and Computer Engineering, University of Mohaghegh Ardabili, Ardabil, Iran*
[b] *School of Technology and Innovations, University of Vaasa, Vaasa, Finland*

PAPER INFO

Paper history:
Received 10 June 2023
Received in revised form 12 September 2023
Accepted 30 September 2023

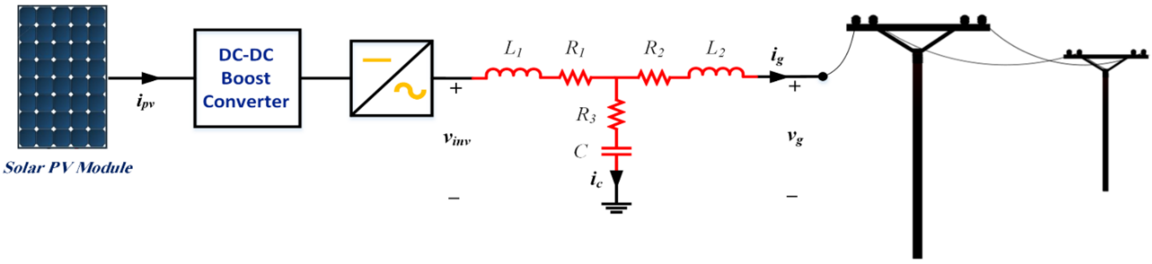
Keywords:
Photovoltaic Array
Grid-Connected Inverter
LCL Filter
Active Damping

ABSTRACT

Grid-connected inverters are considered vital elements for effectively connecting renewable energy sources and distributed generation system applications. Ripple-induced current harmonics in DC link and high switching frequency are the disadvantages of grid-connected inverters that are reduced by LCL filters. However, the intrinsic resonance in the LCL filter leads to instability of the power transmission system. As a result, suitable damping is essential for removing resonance in the LCL filters. The contribution of this paper is to improve the quality of injectable power of LCL filter-based grid-connected photovoltaic array. For this contribution, the stability of the grid-connected inverter has been investigated using active damping method, and maximum power point tracking (MPPT) for the PV array has been performed. The capacitor voltage feedforward active damping method considering computational delay is presented in this paper. By using the inverter-side current feedback beside this method, the proposed control maintains the system's low-frequency specifications independent of the grid impedance changes. It provides high harmonic rejection capability without additional compensators. Also, the number of sensors is decreased due to the alternative measurement of the capacitor voltage instead of the grid voltage for the phase lock loop (PLL). Meanwhile, maximum power point tracking is implemented using the incremental conductance (IC) technique in the boost converter. In addition, a simple and suitable computational method for designing LCL filter parameters is presented, and the system's sensitivity is analyzed. Finally, the simulation has been implemented in MATLAB software that indicates the accurate performance of the control system in injecting the maximum power of the photovoltaic array into the grid and the highly desirable quality of the injectable current to the grid.

doi: *10.5829/ije.2024.37.02b.04*

Graphical Abstract



NOMENCLATURE

P_n	Nominal active power of the system (W)	f_s	Frequency in which Re(A) = 0
P_{pv}	Power of the photovoltaic array (W)	f_{res0}	Resonance frequency of the state in which the grid inductance is zero. (kHz)
v_g	Grid RMS line voltage (V)	f_{swR}	Actual value of switching frequency (kHz)

---

*Corresponding Author Email: *hoseinpour.majid@uma.ac.ir* (M. Hosseinpour)

Please cite this article as: Hosseinpour M, Dastgiri A, Shahparasti M. Design and Analysis of a Power Quality Improvement System for Photovoltaic Generation Based on LCL-Type Grid Connected Inverter. International Journal of Engineering, Transactions B: Applications. 2024;37(02):252-67.

v_{pv}	Photovoltaic voltage (V)	f_g	Grid Frequency (Hz)
v_C	Capacitor voltage (V)	T_{sw}	Switching period
v_{L1}	Inveter-side inductor voltage (V)	T_{d1}	Forward path (control) delay time
v_{L2}	Grid-side inductor voltage (V)	T_{d2}	Feedforward path delay time
V_{inv}	Bridge voltage (V)	G_{d1}	Digital controller computation delay fuction
V_{dc}	Input DC voltage (V)	G_{d2}	Feedforward delay function
i_{pv}	Photovoltaic current (A)	G_Z	Zero-order hold (ZOH) function
i_{L1}	Inverter-side current (A)	G_{feed}	Feedforward coefficient
i_{L2}	Grid-side current (A)	λ_1	Forward path delay time to switching period ratio
I_{sat}	Saturation current of filter inductors (A)	λ_2	Feedforward path delay time to switching period ratio
L_1	Inverter-side inductor (μH)	$\varphi_{\_PI}$	PI controller phase angle at resonant frequency
L_2	Grid-side inductor (μH)	$\varphi_{\_delay}$	Total delay phase angle at resonant frequency
L_g	Grid inductor (μH)	ψ_1	Phase before the phase jump at resonant frequency
L_{20}	Designed inductor (μH)	ψ_2	Phase after the phase jump at resonant frequency
L_{2-C}	Grid-side inductor that can lead to a new resonance	K_{pwm}	Inverter gain
L_R	Actual value of inverter-side inductor	H_F	$G_{feed} * K_{pwm}$
C	Filter capacitor (μF)	C_e	Constant value
f_{sw}	Switching frequency (kHz)	δ	Harmonic attenuation rate
f_{res}	Resonance frequency (kHz)	m	Switching frequency to resonance frequency ratio
f_c	Control frequency (kHz)	k	Integer

1. INTRODUCTION

The microgrid (MG) concept arises from the use of distributed power generators and storage systems to meet demand for optimal operation, voltage/current control, and power quality (1). In the past decade, extensive installation of grid-connected renewable energy sources (RESs) increased including wind energy, solar photovoltaic energy (PV), biomass energy, tidal energy etc. (2). It is noteworthy that solar energy is installed faster and plays a vital role in electricity generation in some countries.

The photovoltaic array is a set of photovoltaic modules, composed of several interconnected cells. PV cells absorb the sun's radiant energy, leading to the current flow between the two layers with the opposite load. Due to the relatively low power generation of solar cells, cells, modules, and arrays can be connected in series or parallel, or in combination to achieve the desired peak output voltage. Electric vehicles, indoor and outdoor lighting, and water pumping can be powered by photovoltaic systems. Grid-connected applications include hybrid systems, power plants etc. (3).

Technological advances and increased production scale leads to continually decreased cost of the PV. PV systems are, however subjected to two fundamental problems, namely a low efficiency in converting solar energy into electrical energy, and a nonlinear design that results in a change in electrical energy produced based on temperature and sunlight (4).

The I-V and P-V curves of solar arrays reveal that to be the most efficient and produce maximum output power, a PV system can only operate at one maximum power point. Various algorithms are available for tracking the maximum power point (MPPT) position (5). It is possible to use several MPPT techniques, including perturbation and observation technique (P&O), incremental conductance technique (IC), ripple correlation technique, short circuit current technique (SCC), open circuit voltage technique (OCV) etc. There have been several reviews of these methods. The methods vary in complexity, cost, convergence speed, sensors required, hardware implementation, and effectiveness.

Inverter is the basic power converter for integrating solar photovoltaic sources with the grid, which converts DC power of the PV array into AC power for injection into the grid. So, the voltage source inverter (VSI) or current source inverter (CSI) is usually combined with DC/DC converter. However, a complex control structure is also needed to achieve better performance and optimal system output. The inverter plays a vital role in controlling the current injection into the grid. Hence, it maintains the voltage value of the DC link at an optimum level and controls the flow of active and reactive power to the grid (6). For this purpose, transformer-less grid-connected photovoltaic inverters are used due to their small size/weight, reduced cost, and high efficiency, which are preferable to inverters with transformers (7). To ensure that grid-connected photovoltaic inverters comply with applicable laws and grid standards, grid-connected photovoltaic inverters should comply with grid standards (8, 9). Several electrical factors should be considered, including leakage current protection, grid frequency monitoring, active and reactive power control, and power quality.

Following IEEE 1547 standard, a filter is required to have a standard harmonic in the inverter output (10). The inductive filter L is one of the most straightforward filters leading to a simple controller design and practical implementation. In grid-connected inverters, LCL filters are preferred due to their high power density, low inductance, and better attenuation of switching frequency harmonics (11). The LCL filter has resonance despite having significant advantages that should be weakened by damping methods to maintain the stability of the inverter (12).

Researchers have proposed a wide range of approaches to address the concerns regarding the stability and injection quality of the LCL-based grid inverter (13). Adding resistance in series or parallel to the inductor or capacitor of the filter is a passive damping method that leads to reduced efficiency despite simplicity (14). Active damping is another method that attracted more attention due to low power loss. However, the high cost due to the number of elements added and the control complexity are the disadvantages of this damping method. The two main active damping methods are single and multi-loop damping methods. Each of these loops includes the capacitor voltage loop (15), the capacitor current loop (16-18), and other loops. It should be noted that single-loop control strategies are advantageous because they are simple and perform satisfactorily due to the absence of supplementary measurement methods or passive damping measures. Additionally, grid-connected inverters are typically controlled by two feedback nodes: a current feedback node on the grid-side inductor (grid-side current feedback) and a current feedback node on the inverter-side inductor (inverter-side current feedback).

A robust quasi-predictive control of LCL-filtered grid converters is applied (19). The proposed method provided the stability of the system against grid impedance changes. A control strategy with inverter-side current feedback-based active damping is proposed by Cai et al. (20), which is usually used to analyze the robustness of the control despite the design. An explanation of how to select the type of controller is not provided (20). A dual-current control method was proposed by Guan et al. (21) to improve the attenuation of resonances in grid-connected inverters with LCL filters. However, this strategy requires more sensors, which leads to an increase in the cost and complexity of the system. Based on negative resonance regulation and grid current feedback; Zhou et al. (22) presented a new active damping algorithm. The switching sensitivity of this algorithm is lower, and the control delay compensation is higher. Using a combination of capacitor current feedback and point of common coupling voltage feedforward, a hybrid active damping algorithm is presented by He et al. (23). The association between these two damping methods ensures the robustness of the active damping in case of parameter change and non-ideal grid conditions. A grid-connected inverter control based on capacitor voltage damping using parallel feedforward compensation was presented by Faiz et al. (24), which uses an alternative feedforward path containing a second-order transfer function to improve system performance against grid impedance changes. Despite the proper stability of the system under investigation, the quality of the injected current to the grid has not been improved in an optimum way. A capacitor voltage feedforward control mechanism is used as part of the proposed approach by Li et al. (25), fully compensating for delays. In cases where the frequency of sampling is insufficiently high, or the analog-to-digital converter boasts high speed, this approach can be used to detect a sampling interval that is due to the computational delay. The effect of capacitor voltage feedforward was analyzed by Zou et al. (26) by considering variable computational delay through zero-pole graphs. Although the forward and feedforward path delays are considered equal despite being independent, in this case, accurate analysis is required. Thus, an accurate way to eliminate the resonance was proposed by Liu et al. (27), where the voltage of the capacitor is on the path of the feedforward. However, this method requires an accurate computational delay in the feedback loop, which creates a new resonance with the existence of an inappropriate delay. Capacitor voltage feedforward damping was applied by Figueredo and Matakas (28), in which the common-mode inductor is added towards AC/DC inverter to reduce the common-mode noise current. However, significant suppression would require a sizeable common-mode inductor, which violates the proposed method. A comparison between the capacitor voltage feedforward active damping and the notch filter for LCL-based grid-connected inverters with inverter-side current was proposed by Rodriguez-Diaz et al. (29). The preference for the capacitor voltage feedforward damping method is observable in the results regarding robustness in nominal conditions, stability in the weak grid, and desirable dynamic response.

A single-loop current control strategy based on active damping and an easy-to-use controller is presented in this paper, which has been applied to various industrial applications as well. Using inverter-side current control, the system is stable with a small computational delay. However, it becomes unstable with the increased computational delay in a limited period. By increasing computational delay, stability is again achieved. Unfortunately, the stable region with a limited computational delay is considerably limited and decreases with increasing resonance frequency. Obviously, in this case, a high-speed processor is needed to detect the small delay time. Also, the delay is too short to fulfill the control codes for high frequency systems properly. Conversely, the region of stability showing enhanced computational delay is highly sensitive to changes in filter parameters. It may experience instability as a result of changes in grid impedance.

Hence, the main contributions of the current paper are as follows:

1. Improving the quality of injectable power of LCL filter-based grid-connected photovoltaic array.
2. Presenting the filter capacitor voltage feedforward with computational delay as a solution to overcome the aforementioned limitations and attenuation of resonance.
3. Developing the stability area of the system.
4. Removing the system sensitivity to grid impedance changes by applying the capacitor voltage feedforward. It is worth noting that the effect of computational delay on the stability of the system has also been investigated in both forward and feedforward paths, and the stability of the system has been significantly improved.
5. Providing high harmonic rejection capability without additional compensators.

Finally, the proposed system is combined with DC/DC boost converter. The injection power quality of the photovoltaic array connected to the LCL filter can be improved using the incremental conductance (IC) technique within the boost converter.

This paper is structured in the following manner; the analysis of the power quality improvement system for LCL-type PV grid-connected inverter is discussed in section two. In section three, an overview of the design and implementation of the LCL filter is provided. In section four, the implementation of the system using MATLAB simulation software is demonstrated and in the last part the paper is concluded with a brief conclusion.

2. SYSTEM ANALYSIS FOR IMPROVING POWER QUALITY OF GRID-CONNECTED PV INVERTER OF THE LCL TYPE:

Figure 1 illustrates the proposed system. The DC/DC boost converter and single-phase inverter based on the LCL filter connect the PV array to the grid.

2. 1. Incremental Conductance (IC) Technique Applied To DC/DC Boost Converter Radiation and temperature changes, shading conditions, and nonlinear output power characteristics lead to fluctuation in the output power of the photovoltaic array. A DC/DC boost converter is proposed in this paper using the IC method to achieve the optimal duty cycle, thus improving injection power quality. MPPT is one of the most essential features of PV power systems. Hence, a DC/DC converter should be considered for MPPT implementation.

At the maximum power point, the slope of the PV array power curve is zero in the IC method $\Delta P/\Delta V = 0$, where $P = VI$. By considering Equation 1, the maximum power point is tracked by comparing the instantaneous conductance I/V with the incremental conductance $\Delta I/\Delta V$. The existing algorithm increases or decreases the reference value until the $\Delta I/\Delta V = -I/V$ condition is met. By achieving the maximum power point, PV array operations remain constant at the optimal point. Good efficiency tracking and automatic voltage regulation of the module without oscillation are the algorithm's advantages.

$$\begin{aligned} &\frac{\Delta I}{\Delta V} = -\frac{I}{V} \;\; if \;\; P = MPP, \\ &\frac{\Delta I}{\Delta V} > -\frac{I}{V} \;\; if \;\; P = MPP, \\ &\frac{\Delta I}{\Delta V} < -\frac{I}{V} \;\; if \;\; P = MPP, \end{aligned} \tag{1}$$

2. 2. System Stability With Inverter-Side Current Control and Feedforward Control of Capacitor Voltage

2. 2. 1. Strategy for Controlling the Inverter-Side Current As shown in Figure 2, the current feedback control strategy for the inverter side is described. Digital controller computational delay indicates by G_{d1}.

$$G_{d1} = e^{-sT_{d1}} \tag{2}$$

The control imposed delay time denoted by the parameter T_{d1}. Furthermore, it is essential to note that the gain of K_{pwm} exhibits a value of V_{dc} when employed in a full-bridge inverter configuration, whereas, in a half-bridge inverter configuration, its gain is equal to $V_{dc}/2$. The mathematical model of the inverter is obtained according to the following Equation:

$$G_{PLANT}(s) = \frac{I_{L1}(s)}{V_{inv}(s)} = \frac{s^2 L_2 C + 1}{s^3 L_1 L_2 C + sL_1 + sL_2} \tag{3}$$

V_{inv} and I_{L1} denote the voltage across the bridge and the current flowing through the inverter side, respectively. Figure 3 illustrates the bode diagram of the G_{PLANT}. The phase undergoes a transition from a value of 90º to -90º at the resonant frequency, as can be seen. The resonant frequency is equal to:

$$f_{res} = \frac{1}{2\pi}\sqrt{\frac{L_1 + L_2}{L_1 L_2 C}} \tag{4}$$

The phase of the system before and after the phase jump at the resonant frequency is equal to:

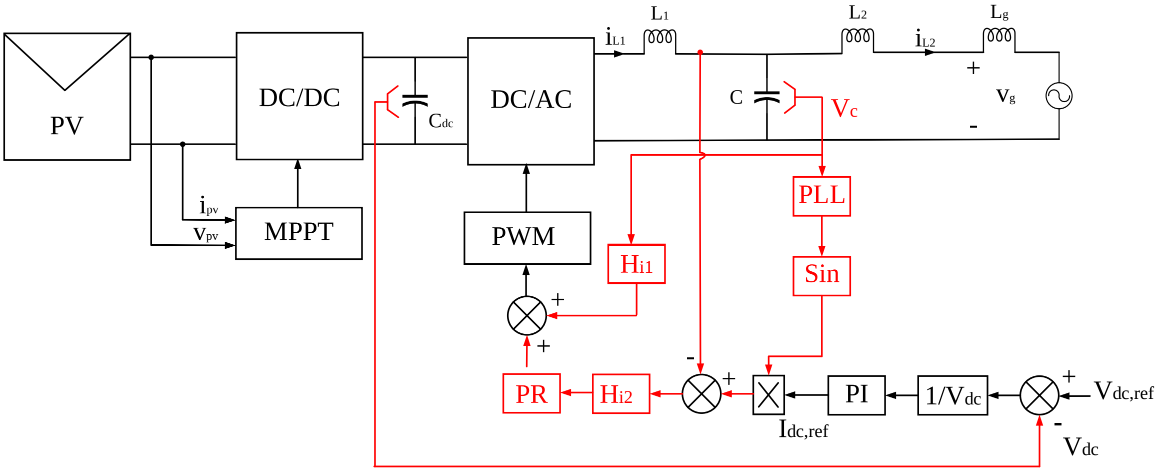


Figure 1. Power circuit scheme of the power quality improvement system for LCL-type grid-connected PV by considering the control system

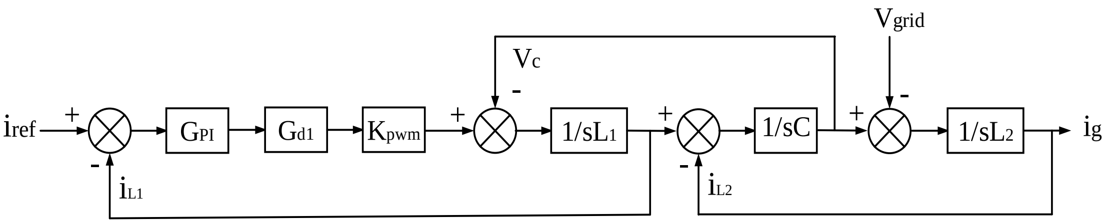


Figure 2. An inverter current feedback control strategy

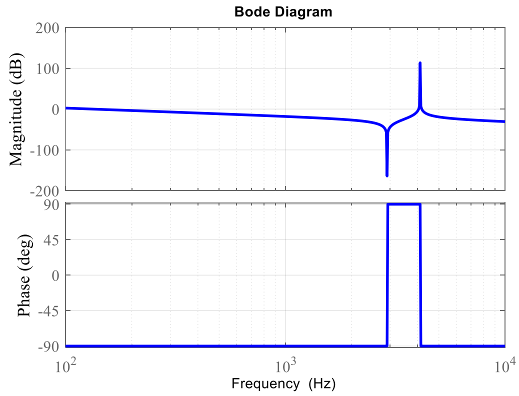


Figure 3. Bode diagram of G_{PLANT}

$$\begin{cases} \psi_1 = 90^{\circ} + \varphi_{\_delay} + \varphi_{\_PI} \\ \psi_2 = -90^{\circ} + \varphi_{\_delay} + \varphi_{\_PI} \end{cases} \tag{5}$$

where the phase angles resulting from the total delay and the PI controller at the resonant frequency are represented by $\varphi_{\_delay}$ and $\varphi_{\_PI}$, respectively. It is important to note that the total delay includes the delay caused by the computation and the zero-order holding element (ZOH). Thus:

$$\varphi_{\_delay} = -f_{res}(T_{d1} + 0.5T_{sw}) \times 360^{\circ} \tag{6}$$

It is worth noting that at high frequencies, the P controller can be used instead of a PI controller. In other words, $\varphi_{\_PI} \approx 0$ at the resonance frequency. If ψ_1 is within the range that appeared in Equation 7, the system is unstable.

$$-180^{\circ} - 360^{\circ} k < \psi_1 < 0^{\circ} - 360^{\circ} k \tag{7}$$

where k is an integer, it is worth mentioning that Equation 7 is a sufficient but unnecessary condition.

In the next step, some unstable areas are calculated. With the definitions of $m = f_{sw}/f_{res}$ and $\lambda_1 = T_{d1}/T_{sw}$ and replacing Equation 5 in Equation 7, the system becomes unstable under the following condition:

$$km + \frac{m-2}{4} < \lambda_1 < km + \frac{3m-2}{4} \tag{8}$$

Unstable areas are changed due to grid inductance changes, resonance frequency, and m. The stable and unstable regions can be seen in Figure 4 as per the formulation presented in Equation 8, and upon contemplation of the variation in the parameter m from a value of 3 to 5, with k set at 0. Therefore, three regions exist. Region I is stable and constrained. Region II is unstable. Region III is also stable if the m value rises to three. However, as the grid inductance increases, the value of *m* increases, and the stability region approaches the unstable boundary.

2. 2. 2. Feedforward Control Strategy Using Capacitor Voltage

In order to maintain the stability of the system when taking into account the variation of the grid impedance, it is essential that the time delay T_{d1} be sufficiently small. In order to resolve this problem, the voltage feedforward method is used to resolve the slight computational delay in the control path. Unfortunately, the voltage feedforward method is impractical or requires a faster CPU. Therefore, the system's stability is then assessed following this method.

Since the capacitor voltage is measured simultaneously with the grid voltage for the Phase Locked Loop (PLL), there is no need to measure the capacitor voltage separately, which gives this method the advantage over other active damping methods. An illustrative depiction of the control methodology the utilization of capacitor voltage feedforward is presented in Figure 5. T_{d1} and T_{d2} are the delays in the forward and feedforward paths, respectively. G_{feed} is also called the feedforward coefficient. The delay relations in both paths are calculated as follows.

$$\begin{cases} G_{d1} = e^{-sT_{d1}} \\ G_Z = \dfrac{1}{T_{sw}}\dfrac{1-e^{-eT_{sw}}}{s} \approx e^{-0.5sT_{sw}} \\ G_{d2} = e^{-sT_{d2}} \end{cases} \tag{9}$$

By using capacitor voltage feedforward, the open loop transfer function of the system is calculated and can be seen in Equation 10:

$$GH(s) = \frac{\left(s^2L_2C+1\right)G_{PI}G_{d1}G_ZK_{pwm}}{s^3L_1L_2C + sL_1 + sL_2 - sL_2G_{d2}G_{feed}G_ZK_{pwm}} \tag{10}$$

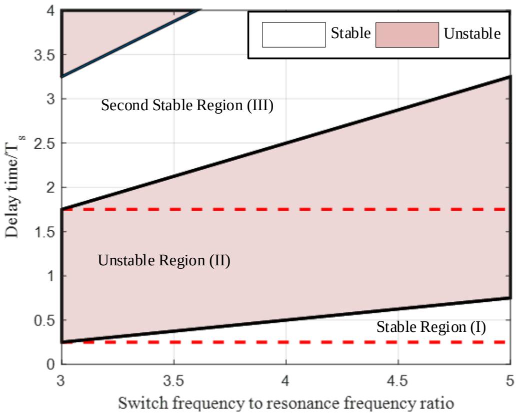


Figure 4. Regions with the varying resonant frequency

In the first step, it is necessary to determine the coefficient of the feedforward path using the characteristic Equation of the open loop transfer function, which is as follows:

$$s(s^2L_1L_2C + L_1 + L_2 - L_2G_{d2}G_{feed}G_zK_{pwm}) = 0 \tag{11}$$

Given that $|G_{d2}G_z| = 1$, the system will have no unstable open-loop pole if the Equation 12 is established.

$$G_{feed}K_{pwm} \le 1 \tag{12}$$

Under the assumption that the unit feedforward is a joint selection as well as the ideal feedforward in the context of the unit feedforward, G_{feed} can be calculated as follows:

$$G_{feed} = \frac{1}{K_{pwm}} \tag{13}$$

A quick check of the control and feedforward paths computation delays is conducted. Since feedforward control requires solely one additional operation, a very small value of T_{d2} can be chosen. However, T_{d1} caused by complex control codes cannot be as small as the computational delay T_{d2}. It is worth noting that the control frequency (f_c) can be equal to or twice the switching frequency (f_{sw}), which is considered equivalent in this article.

After examining the delay in the feedforward path by considering the non-ideal conditions, the G_{PLANT} function is changed to the modified function according to Equation 14.

$$G_P(s) = \frac{s^2L_2C+1}{s^3L_1L_2C + sL_1 + sL_2 - sL_2G_{d2}G_{feed}G_zK_{pwm}} \tag{14}$$

As a first step, the unit delay condition is analyzed in order to demonstrate the effect of the feedforward control strategy by taking into account the delay, where:

$$T_{d2} = T_{sw} \tag{15}$$

The bode diagram of $G_P(s)$ by considering the various resonance frequencies is shown in Figure 6. Based on the application of capacitor voltage feedforward, it is possible to summarize the different results in the following manner.

- Larger domain in low frequency
- Phase drop at low frequency without considering feedforward
- Elimination of the main positive resonance (However, producing a new resonance is possible whose phase drops by 180° around this resonance (according to Figure 6(b))

As a result, there is no significant difference between the feedforward path and without the feedforward path for the $G_P(s)$ function to distinguish between these two paths.

By defining $G_{feed} * K_{pwm} = H_F$ where single feedforward and without feedforward identified by $H_F = 1$ and $H_F = 0$, respectively, the amplitude-frequency function of $G_P(s)$ can express according to Equation 16 where $\lambda_2 = T_{d2}/T_{sw}$.

$$G_{P\_MF} = \frac{1-\omega^2 L_2 C}{-j\omega^3 L_1 L_2 C + j\omega L_1 + j\omega L_2 - jH_F \omega L_2 e^{-j(\lambda_2+0.5)\omega T_{sw}}} = \frac{1-\omega^2 L_2 C}{\omega}\frac{1}{A} \tag{16}$$

Considering the unit delay, $\lambda_2 = 1$.

$$\begin{aligned} A &= -H_F L_2 \sin 1.5\omega T_{sw} \\ &+ j(L_1 + L_2 - \omega^2 L_1 L_2 C - H_F L_2 \cos 1.5\omega T_{sw}) \end{aligned} \tag{17}$$

$L_1 + L_2 - \omega^2_{res} L_1 L_{2} 0 C = 0$ at the resonance frequency. G_P domain in f_{res} is equal to :

$$|G_P| = \frac{1}{H_F \omega_{res} L_1} \tag{18}$$

Therefore, there is an infinite resonance in the value of $H_F = 0$, which will be removed if the two conditions $H_F = 1$ and $\omega_{res} L_1 > 1$ are established. As can be seen, the G_P domain in f_{res} has a reverse relation with H_F, ω_{res}, and L_1. However, it is possible to create a new resonance at this time.

The condition that limits this, is the equality of the real and imaginary part of the expression A with zero value at the same frequency.

$$\mathrm{Re}(A) = \mathrm{Im}(A) = 0 \tag{19}$$

$Re(A) = 0$ occurs when

$$f_s = \frac{kf_{sw}}{3}, \quad k = 0,1,2,... \tag{20}$$

f_s is the frequency that causes the real value of A to become zero. Also, $k = 0$ is meaningless due to the representation of the dc signal and does not affect the infinite gain in the dc signal system stability. It is worth mentioning that the possibility of creating an infinite resonance peak is provided by $Im(A)$ becoming zero at f_s.

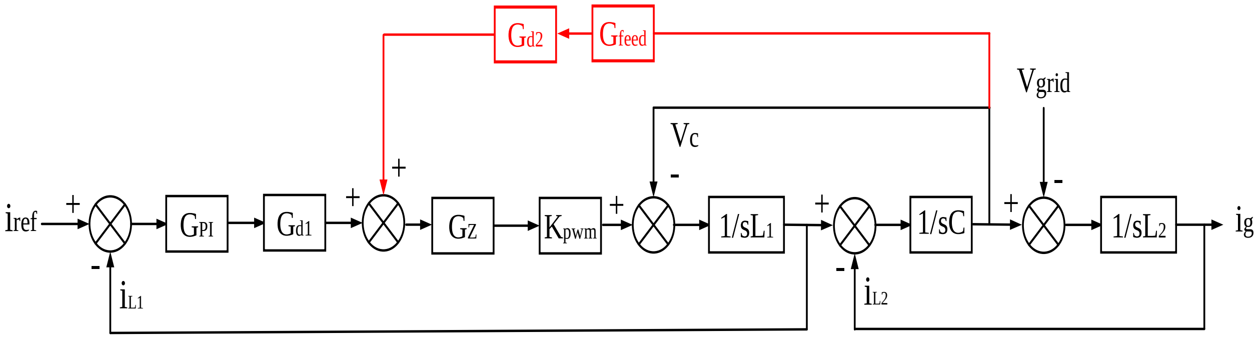
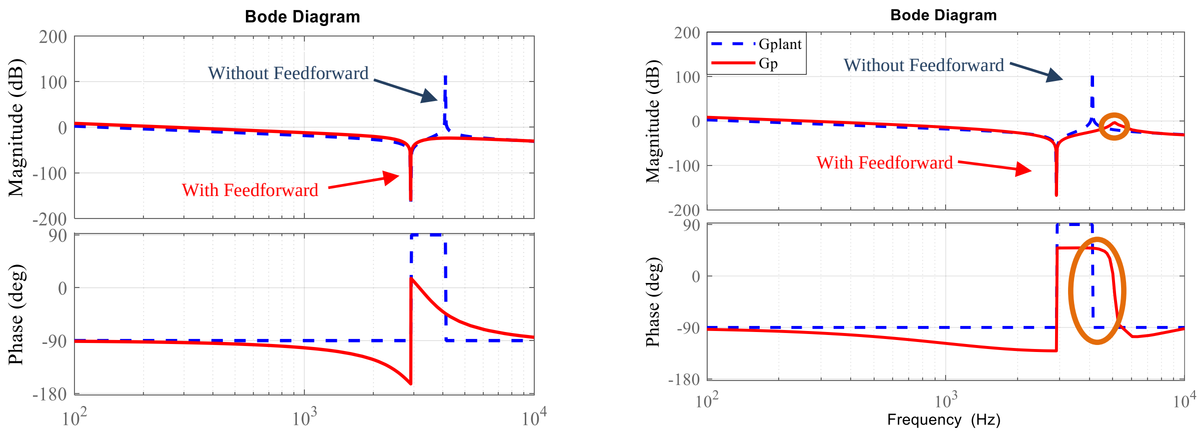


Figure 5. Capacitor voltage feedforward control strategy



(a)

(b)

Figure 6. $G_P(s)$ bode Diagram with various resonance frequencies (a) Without resonance status (b) With new resonance status

In addition, the finite resonance peak is created with low *Im(A)* in the f_s. Similarly, if the *Im(A)* value in the f_s is large enough, no new resonance peak will occur. However, if Equation 20 is replaced by Equation 17, then a new resonance peak will occur as follows:

$$L_1 + L_2 - 4\pi^2 L_1 L_2 C \frac{k^2 f_{sw}^2}{9} - (-1)^k L_2 = 0 \tag{21}$$

By solving Equation 21, a new component for the grid side inductor can define (L_{2-C}), which leads to a new resonance. Hence, a new infinite resonance will exist, by equalization of the LCL filter grid-side inductor with the following value.

$$L_{2-C} = \frac{L_1}{4\pi^2 L_1 C \frac{k^2 f_{sw}^2}{9} + (-1)^k - 1} \tag{22}$$

It is worth noting that a smaller value of the L_{2-C} than the value of the grid-side inductor L_2 will prevent the existence of a new infinite resonance. So, an inverter with zero grid equivalent inductor has the minimum grid-side inductor, theoretically L_{20}. On the other hand, the $k = 1$ means a maximum value of L_{2-C}. Therefore, only $k = 1$ state is considered.

Finally, it can be concluded that there will be an infinite resonance if L_{2-C} is greater than L_{20}. Therefore, there will be a new resonance if

$$L_{2-C} = \frac{L_1}{4\pi^2 L_1 C \frac{f_{sw}^2}{9} - 2} > L_{20} \tag{23}$$

The $L_1 + L_{20} - 4\pi^2 f^2_{res0} L_1 L_{20} C = 0$ can now be placed in Equation 23 and resolved that the final condition is obtained according to the following Equation 24.

$$L_{20} < \frac{L_1\left(\frac{f_{sw}^2}{9 f_{res0}^2} - 1\right)}{\left(2 - \frac{f_{sw}^2}{9 f_{res0}^2}\right)} \tag{24}$$

where f_{res0} is the resonant frequency of the state in which the grid inductance is zero. According to Equation 24, when f_{res0} is between the values of $0.235 f_{sw}$ and $0.333 f_{sw}$, L_{20} has a positive value. Also, a new infinite resonance can occur by adopting the unit feedforward.

The infinite resonance peak occurs with the equality of ($L_2 = L_{2-C}$), which can be seen in the following Equation.

$$L_{2-C} = L_2 = \frac{L_1\left(\frac{f_{sw}^2}{9 f_{res}^2} - 1\right)}{\left(2 - \frac{f_{sw}^2}{9 f_{res}^2}\right)} \tag{25}$$

However, such a state is extraordinary and difficult to establish. Therefore, the resonance peak value is essential when L_2 is around L_{2-C}. The gain is shown by calculating the G_P amplitude when the grid inductance changes around L_{2-C}.

A variable describing the ratio between the resonance frequency and the switching frequency (f_{res}/f_{sw}) and the (L_2/L_1) ratio is utilized in arrange to maintain a strategic distance from employing a particular illustration to demonstrate the alteration in resonance phenomenon. The amplitude is calculated according to Equation 26. It can be seen that there is a term of ($f_{sw}L_1$) in Equation 26, which is a constant term and has no effect on the inclination of resonance peak changes.

$$|G_P| = \frac{1}{f_{sw} L_1} \times \left| \frac{1 - \left(1 + \frac{L_2}{L_1}\right) \frac{f^2}{f_{sw}^2} \frac{f_{sw}^2}{f_{res}^2}}{2\pi \frac{f}{f_{sw}}} \right| \times \frac{1}{\sqrt{\left(\frac{L_2}{L_1} \sin \frac{3\pi f}{f_{sw}}\right)^2 + \left(\left(1 + \frac{L_2}{L_1}\right)\left(1 - \frac{f^2}{f_{sw}^2} \frac{f_{sw}^2}{f_{res}^2}\right) - \frac{L_2}{L_1} \cos \frac{3\pi f}{f_{sw}}\right)^2}} \tag{26}$$

G_P resonance peak is analyzed assuming $f_{sw} = 10kHz$ and $L_1 = 1mH$. The new resonance peak can be accomplished by means of choosing different values for the actual value of the inverter side inductor and the switching frequency (L_R, f_{swR}), and adding a constant value of C_e. The constant value of C_e is equal to

$$C_e = 20 \times \lg \frac{f_{sw} L_1}{f_{swR} L_R} \tag{27}$$

G_P resonance peak can be calculated through MATLAB software considering specific values for (f_{res}/f_{sw}) and (L_2/L_1). It is worth noting that λ_2 is considered equal to one in all the above relations, which can be assumed as a variable. By reducing the computational delay λ_2 and reducing the maximum value of L_{2-C}, the limitation of the grid-side inductor is also reduced.

Next, the design of the LCL filter is analyzed, and appropriate values are obtained for each element.

3. DESIGN OF LCL FILTER

The grid-side inductor includes the equivalent grid inductor and designed inductor under the following relation in the single-phase inverter system described in Figure 1.

$$L_2 = L_g + L_{20} \tag{28}$$

L_g is the grid equivalent inductor, and L_{20} is the designed inductor. L_1 and C, shown in Figure 1, are also called inverter-side inductor and filter capacitor, respectively.

It attempts to meet the grid code requirements by attenuating high-order harmonics on the grid side by using LCL filters (30). The following input data are required to design the LCL filters:

- Grid RMS line voltage V_g
- Nominal active power of the system P_n
- Nominal frequency of grid voltage f_g
- Converter switching frequency f_{sw}
- Saturation current of filter inductors I_{sat}

Then the LCL filter parameters are adjusted according to the following steps.

3. 1. Total Inductor Maximum Value

To prevent losses and filter voltage drops, it is recommended to consider the sum of LCL filter inductor values as small as possible. The total inductance value should be less than 0.1 per unit to improve the system's speed and dynamics.

$$L_{T\max} = \left(L_1 + L_2\right)_{\max} = 10\% L_{TBase} \tag{29}$$

$$\begin{cases} L_{TBase} = \dfrac{Z_{Base}}{2\pi f_g} \\ Z_{Base} = \dfrac{V_g^2}{P_n} \end{cases} \tag{30}$$

$$L_{T\max} = 10\% \frac{V_g^2}{2\pi f_g P_n} \tag{31}$$

Here, L_{TBase} is the base value of the total inductor, and Z_{Base} is the base impedance.

3. 2. LCL Filter Capacitor With a Maximum Value

The power consumed by the capacitor Q_C should be 2% to 5% of the nominal power injected into the grid. For this purpose, the selection range of the capacitor is determined as follows:

$$\begin{cases} C_f \le \dfrac{5\% P_n}{2\pi f_g V_g^2} \\ C_f \ge \dfrac{2\% P_n}{2\pi f_g V_g^2} \end{cases} \tag{32}$$

An additional point to note is that when the capacitor amount is low, it is recommended to select very high inductor values. To achieve satisfactory results, it is recommended to start with a capacitor value equal to half of the maximum value and increase it to the maximum value if it is unsatisfied.

3. 3. Adjustment of Inverter Side Inductor

In the most severe case where the ripple current originates from the pulse voltage produced by the converter, the inductor on the inverter side of the converter is designed to reduce the ripple current.

$$\begin{aligned} & V_{i\max} = \sqrt{V_{g\max}^2 + \left(L_{T\max}\left(2\pi f_g\right) I_{2\max}\right)^2} \\ & V_{dc\min} = \sqrt{2} V_{i\max} \\ & I_{1\max} = \frac{P_n}{V_g} \\ & I_{sat} = 20\% I_{2\max} \end{aligned} \tag{33}$$

Finally, L_{imin} is equal to

$$L_{1\min} > \frac{V_{dc}}{16 f_{sw}\left(I_{sat} - I_{1\max}\right)} \tag{34}$$

It is noteworthy that the value of L_{imin} should be smaller than the value of L_{Tmax}.

$$L_{1\min} \le L_{T\max} \tag{35}$$

3. 4. Adjustment of Grid Side Inductor

Depending on a grid code standard and requirements, the grid side inductor should limit grid current harmonics compared to the IEEE519-1992 standard. This provides a total harmonic disturbance (THD) limit of less than 5% for grid currents.

Following is a formula that calculates the relation between the inductor on the inverter side and the inductor on the grid side.

$$L_2 = aL_1 \tag{36}$$

where

$$0 \le a \le a_{\max} \rightarrow a_{\max} = \frac{L_{T\max}}{L_1} - 1 \tag{37}$$

The harmonic attenuation rate δ is used to determine the coefficient a, which represents the relation between the currents on the inverter and grid sides at the switching frequency. The relation between these two coefficients is:

$$a = \frac{1+\delta}{\delta a_1} \rightarrow a_1 = L_1 C_f \omega_{sw}^2 - 1 \tag{38}$$

A proper design should take into account the following conditions for harmonic attenuation rate:

$$\begin{cases} \delta > \dfrac{36L_1 - (2\pi f_{sw} L_1)^2 C_{f\max}}{a_3 (2\pi f_{sw})^2 - 36a_2} \Rightarrow 1^{st}\ \delta\ condition \\ \delta < \dfrac{4L_1 - (2\pi f_{sw} L_1)^2 C_{f\min}}{b_3 (2\pi f_{sw})^2 - 4b_2} \Rightarrow 2^{nd}\ \delta\ condition \end{cases} \tag{39}$$

The unknown coefficients of Equation 39 are calculated according to the following relations:

$$\begin{cases} a_1 = L_1 C_f \left(2\pi f_{sw}\right)^2 - 1 \\ a_2 = L_1 + a_1 L_{g\,\max} + a_1 L_1 \\ a_3 = \left(L_1 + a_1 L_{g\,\max}\right) L_1 C_{f\,\max} \end{cases} \quad (40)$$

$$\begin{cases} b_2 = L_1 + a_1 L_{g\,\min} + a_1 L_1 \\ b_3 = \left(L_1 + a_1 L_{g\,\min}\right) L_1 C_{f\,\min} \end{cases}$$

In addition, to ensure the change of resonance frequency in the stability region, the desired harmonic attenuation rate must be greater than the minimum harmonic attenuation rate δ_{min}, which corresponds to a_{max}.

$$\delta > \delta_{\min} = \frac{1}{\left|1 + a_{\max} a_1\right|} \Rightarrow 3rd\,\delta\,condition \quad (41)$$

It is worth noting that the current harmonics will decrease by decreasing the value of δ. The harmonic disturbance of the total grid current is also reduced by reducing the harmonic attenuation rate of δ. The desired harmonic attenuation rate may be easily determined by applying Equations 39-41. As a next step, the value of coefficient a is determined by utilizing Equation 38, and finally, the value of the grid side inductor is determined by utilizing Equation 36.

$$f_{res} = \frac{1}{2\pi}\sqrt{\frac{L_1 + L_2}{L_1 L_2 C_f}}$$
$$10 f_g \le \frac{f_{sw}}{6} \le f_{res} \le 0.5 f_{sw} \quad (42)$$

4. SIMULATION RESULTS

In this section, to verify the performance of the inverter side inductor current feedback technique and the proposed filter capacitor voltage feedforward, the PV array system grid-connected single-phase inverter is simulated in MATLAB software, and the results are discussed in detail. Tables 1 and 2 present a list of parameters used for each system.

The stability of the proposed system has been investigated under five different scenarios, which are:

- Scenario 1: Stable condition
- Scenario 2: Sudden changes in solar radiation
- Scenario 3: Grid voltage drop condition (sag)
- Scenario 4: Grid voltage increase condition (swell)
- Scenario 5: Applying the fifth and seventh harmonics

Figure 7 shows the I-V and P-V curves of the SunPower SPR-E20-327 solar array for two different radiation intensities, 1000 W/m^2, and 800 W/m^2, respectively. In the radiation of 1000 W/m^2, the current of the solar array is equal to 94.17 A, and the voltage is equal to 2.328 V. It is worth mentioning that the maximum power of the PV array is equal to 5888 W. Now different scenarios will be examined.

TABLE 1. Parameters used in the simulation

Parameter	Symbol	Value
Grid Voltage and Frequency	(V_g, f_g)	230 V, 50 Hz
Boost Switching Frequency	($f_{sw,B}$)	5 kHz
Inverter Switching Frequency	(f_{sw})	10 kHz
Boost Inductor	(L_B)	5 mH
DC-Link Capacitor	(C_{dc})	6 μF
Inverter Side Inductor	(L_1)	600 μH
Grid Side Inductor	(L_2)	265 μH
Grid Equivalent Inductor	(L_g)	550 μH
Filter Capacitor	(C)	8 μF
Controller Parameters	(k_P, k_i)	1, 4500
Capacitor Voltage Feedforward Coefficient	(H_{i1})	0.005
Inverter Side Current Inductor Feedback Coefficient	(H_{i2})	0.15

TABLE 1. PV Panel Parameters (With 25°C Temperature and 1000 W/m^2 Radiation Intensity)

Parameter	Symbol	Value
Open Circuit Voltage	(v_{oc})	65.1 V
Maximum Voltage	(v_{mp})	54.7 V
Short Circuit Current	(i_{sc})	6.46 A
Maximum Current	(i_{mp})	5.98 A
Maximum Power	(P_m)	327.106 W
Number of Series Panels	(N_s)	6
Number of Parallel Panels	(N_p)	3

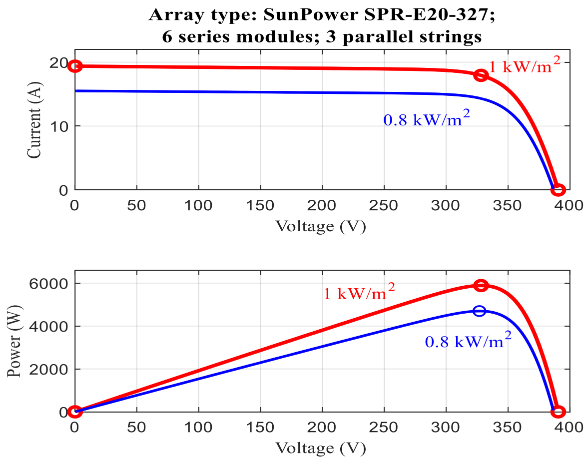
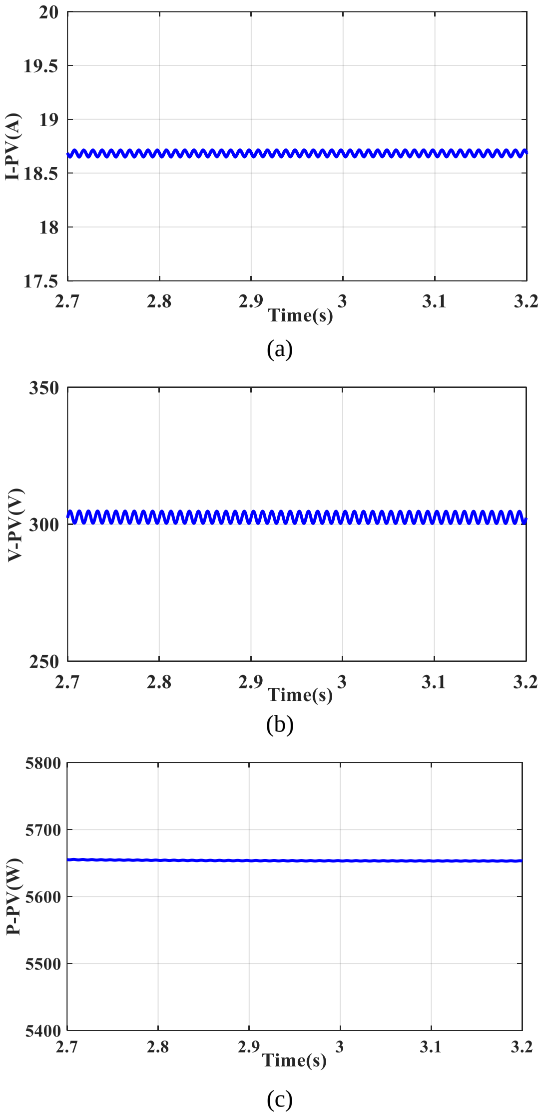


Figure 7. I-V and P-V curve of the solar array

The four mentioned scenarios define different disturbances caused by PV and single-phase grid. During these scenarios, dynamic performance such as grid voltage V_g, grid current I_{L2}, grid power P_g, and grid current THD are evaluated to check grid condition. Also, the dynamic behavior of v_{pv}, i_{pv}, P_{pv}, and V_{dc} are investigated to analyze the PV performance during different test scenarios.

Scenario 1: In Figures 8-10, the dynamic behavior of the PV system connected to the grid with a single-phase inverter is investigated in stable conditions.

PV current and voltage are depicted in Figures 8(a) and 8(b). The current and voltage equal 18.7 A and 305 V, respectively. The maximum power of the solar array can also be seen in Figure 8(c), which is equal to 5620 W. Notably, the simulation results have little difference from the expected values. The DC link capacitor voltage of the boost converter output is represented in Figure 9, which tracks the reference voltage $V_{dc} = 500\ V$ very well. The waveform of the current injected into the grid can also be seen in Figure 10(a), which shows the injection of sinusoidal current into the grid well. Figure 10(b) shows the grid voltage and current waveforms. The grid voltage and current waveforms are completely sinusoidal, with no phase difference between voltage and current. The injected power to the grid is also shown in Figure 10(c), which is equal to 5620 W according to the PV array production power. Finally, the grid current THD is shown in Figure 11, equal to 0.78%.



(a)

(b)

(c)

Figure 8. PV output graphs under stable conditions. a) PV output current b) PV output voltage c) PV output power

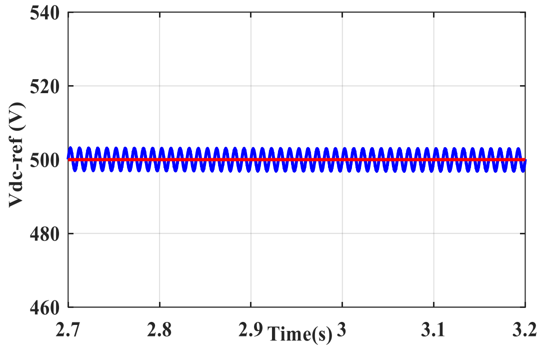
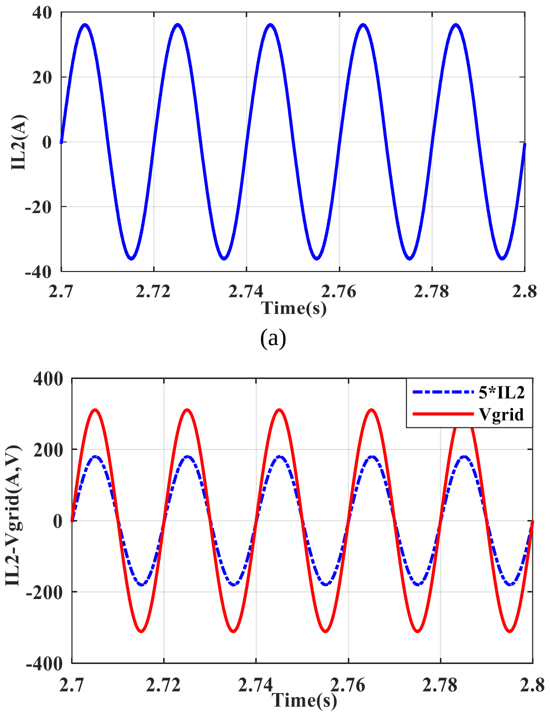


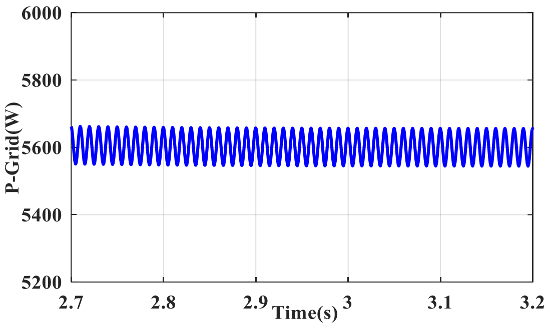
Figure 9. DC-Link capacitor voltage under stable condition

Scenario 2: In Figures 12-15, the dynamic behavior of the PV system connected to the grid with a single-



(a)

(b)



(c)

Figure 10. Dynamic performance of PV power injection system in stable condition. a) grid current b) grid current and voltage c) grid power

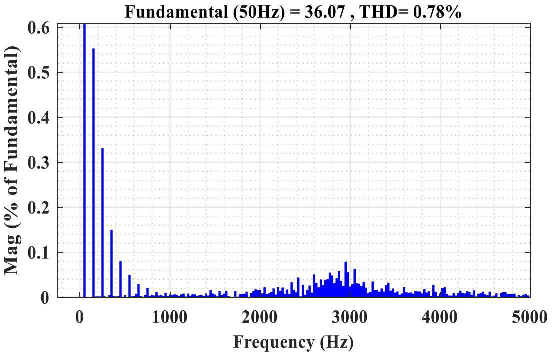


Figure 11. Grid current THD under stable condition

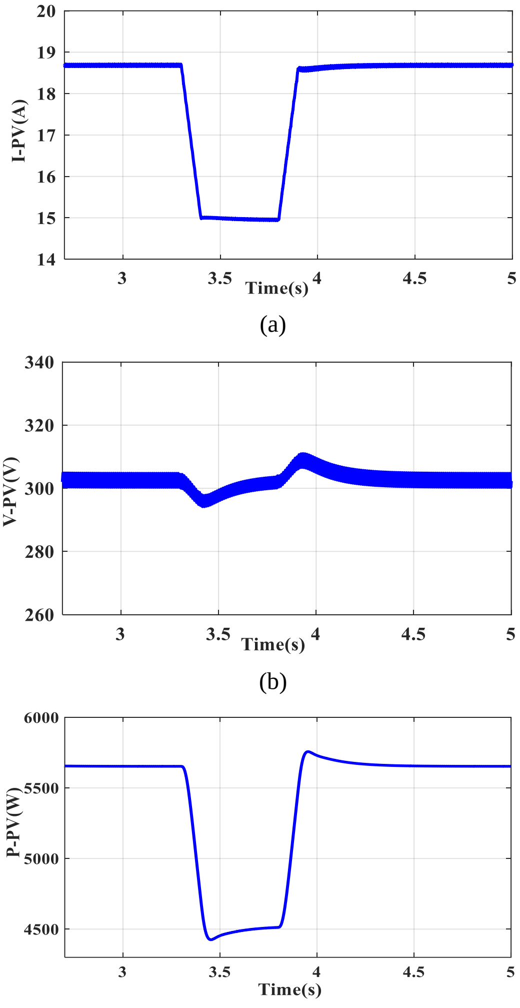
phase inverter was investigated in the solar radiation intensity sudden change condition. The radiation intensity decreases from 1000 W/m^2 to 800 W/m^2 in 3.3-3.9 seconds. In the 800 W/m^2 radiation, the current of the solar array is equal to 14.37 A, and the voltage is equal to 326.9 V. It is worth mentioning that the maximum power of the PV array is equal to 4698W.

The PV current increases or decreases depending on sudden changes in solar radiation, as shown in Figure 12(a). As shown in Figure 12(b), a slight change in PV voltage occurs during this scenario to achieve maximum power. The power produced by the solar array during the decreasing radiation intensity is also shown in Figure 12(c), which decreases from 5620 watts to 4450 watts with a slight difference compared to the P-V curve in Figure 7 and then increases.

The slight difference between the obtained results with the values of the curves in Figure 7 confirms the accuracy of the simulation. The DC link capacitor voltage can be seen in Figure 13. Due to the effect of the incremental conductance technique and the applying MPPT, the DC link voltage adaptively maintains a constant value of 500 V. The grid current THD is shown in Figure 14, whose value equals 1.02%.

Figure 15(a) shows the injection of current into the grid. There is a corresponding decrease in the injection current to the grid during the radiation reduction period. A return to its initial level follows this point. Figure 15(b) shows the injection current in phase with the grid voltage while the radiation reduction period. The grid power during decreasing radiation intensity is also shown in Figure 15(c). The change in the power value is according to the production power of the PV array.

Scenario 3: In Figure 16, the dynamic behavior of the PV system connected to the grid with a single-phase inverter under grid voltage drop (sag) is examined during 3.1-3.35 seconds when the grid voltage is reduced by 20%. The grid current increment under the grid voltage reduction, the in-phase current and grid voltage, and the sinusoidal shape of the grid current can be seen in Figure 16.



(a)

(b)

(c)

Figure 12. PV output graphs under sun radiation intensity sudden change condition. a) PV output current b) PV output voltage c) PV output power

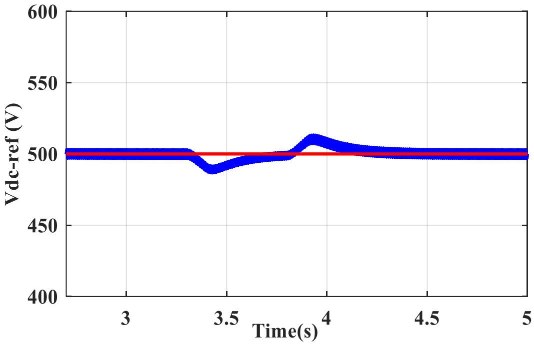


Figure 13. DC-Link capacitor voltage under sun radiation intensity sudden change condition

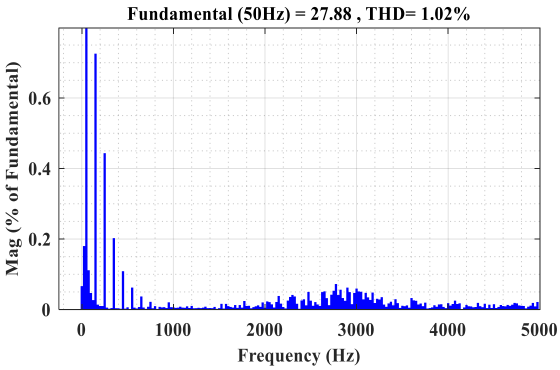
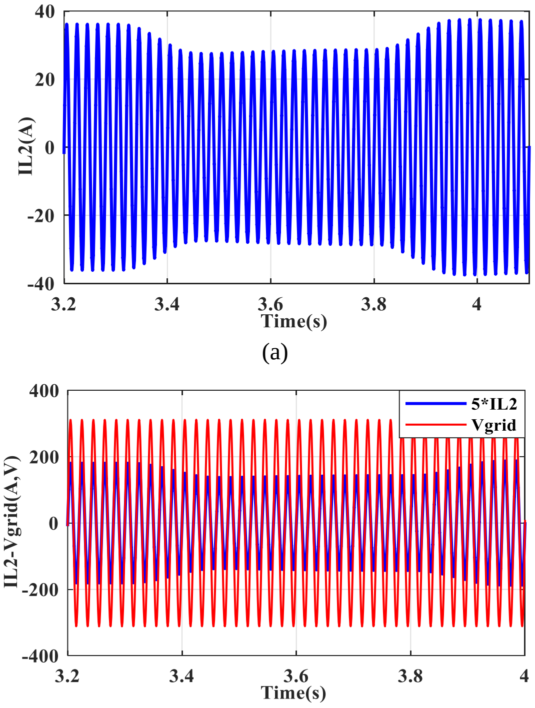
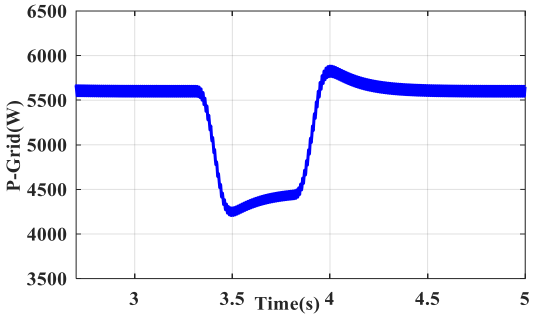


Figure 14. Grid current THD under sun radiation intensity sudden change condition



(a)

(b)



(c)

Figure 15. Dynamic performance of PV power injection system under sun radiation intensity sudden change condition. a) grid current b) grid current and voltage c) grid power

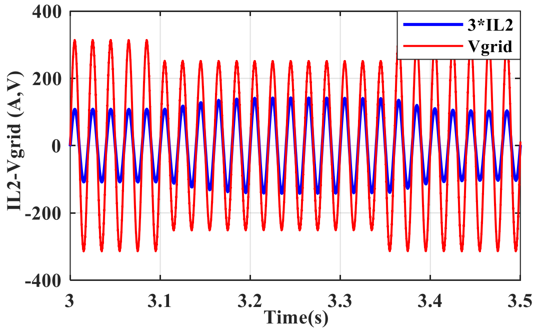


Figure 16. Dynamic performance of PV power injection system under grid voltage drop (sag) condition. Grid current and voltage

Scenario 4: In Figure 17, the dynamic behavior of the PV system connected to the grid with a single-phase inverter under the condition of the grid voltage increment (swell) in the period 3.10-3.35 seconds is examined when the grid voltage increases by 10%. The grid current drop under the rise of grid voltage, and the in-phase current and grid voltage can be seen in Figure 17.

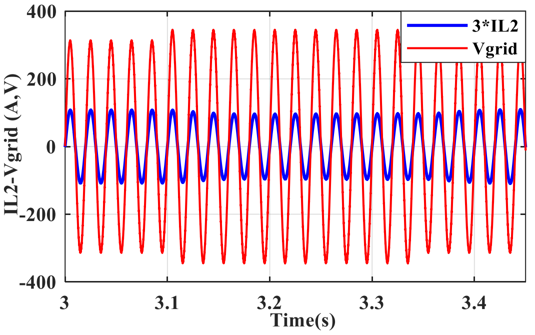


Figure 17. Dynamic performance of PV power injection system under grid voltage increment (swell) condition. Grid current and voltage

Scenario 5: A single-phase inverter is used in this scenario to connect a PV system to the grid, and fifth and seventh harmonics are applied to the system to determine its dynamic behavior. Figure 18 illustrates the grid voltage and current, which have fifth and seventh harmonics, with the in-phase voltage and grid current remaining unchanged. Notably, under grid voltage drop, the THD of the grid current remains below 5% according to the IEEE-519 standard, which can be seen in Figure 19.

Finally, the comparison of the proposed damping method with the current methods that is made in the area of enhancing the performance of the grid-connected inverter control system is presented in Table 3. The compared references are evaluated in terms of control strategy, filter design, and control delay. Moreover, the THD value of the current injected into the grid under the assumed conditions of this paper is also presented in Table 3. It is worth mentioning that the aim of presenting the THD value is only to show a comprehensive comparison. Whereas the different conditions considered in each reference lead to the uncited comparison of the harmonic values with each other. It is obvious that except for data reported by Cai et al. (20), the current injected into the grid using the proposed approach has a lower THD and a higher quality than the others. The number of filters utilized in the grid side indicates that the proposed scheme uses lower number of filters versus most of compared schemes resulting in lower sensor cost.

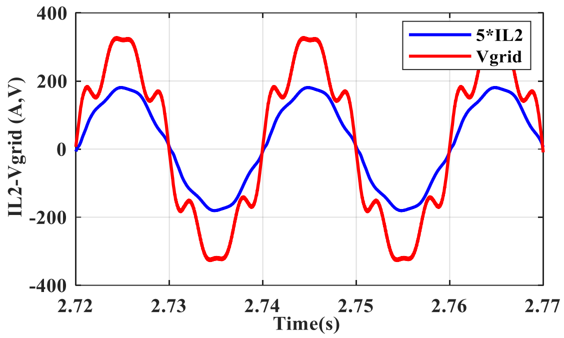


Figure 18. Grid voltage and current under fifth and seventh harmonics

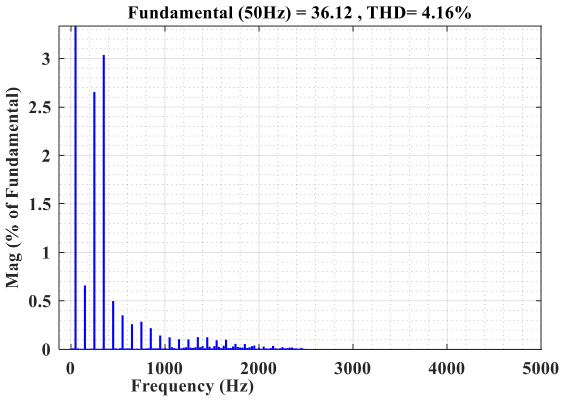


Figure 19. Grid current THD under fifth and seventh harmonics

TABLE 3. Comparison of the proposed method with recent published articles

Ref.	Control Strategy	Filter Design	Grid Side Sensors Count	THD Value
[19]	Capacitor Voltage Feedback	X	3	2.62%
[20]	Inverter Side Current Feedback	X	3	0.38%
[21]	Inverter & Grid Side Current Feedbacks	X	3	3.02%
[22]	Grid Side Current Feedback	X	2	2.3%
[23]	Capacitor Current Feedback & Common Coupling Voltage Feedforward	X	3	2.2%
[24]	Quasi-Predictive Control	X	3	4.35%
Proposed	Capacitor Voltage Feedforward & Inverter Side Current Feedback	☑	2	0.78%

5. CONCLUSION

In this article, a grid-connected PV array system with a LCL filter-based single-phase inverter is proposed. Improving the dynamic performance of the control system without adding power losses and improving the quality of injected power of the photovoltaic array connected to the grid are among the aims of the proposed system.

In this article, the filter capacitor voltage feedforward is used along with the inverter side current feedback, which leads to increasing the stability of the system. In addition, the LCL filter resonance is also damped by the presented control methods. As a result of the MATLAB simulation software, the stability of the system is demonstrated under different scenarios, which confirms the accuracy of the content presented in this article.

6. ACKNOWLEDGMENT

This work has been financially supported by the research deputy of University of Mohaghegh Ardabili. The grant number was 3457.

6. REFERENCES

1. Mobashsher MM, Keypour R, Savaghebi M. Distributed optimal voltage control in islanded microgrids. International Transactions on Electrical Energy Systems. 2021;31(11):e13045. 10.1002/2050-7038.13045
2. Sagar G, Debela T. Implementation of optimal load balancing strategy for hybrid energy management system in dc/ac microgrid with pv and battery storage. International Journal of Engineering, Transactions A: Basics. 2019;32(10):1437-45. 10.5829/ije.2019.32.10a.13
3. Manna S, Singh DK, Akella AK, Kotb H, AboRas KM, Zawbaa HM, et al. Design and implementation of a new adaptive MPPT controller for solar PV systems. Energy Reports. 2023;9:1818-29. 10.1016/j.egyr.2022.12.152
4. Tang C-Y, Wu H-J, Liao C-Y, Wu H-H. An optimal frequency-modulated hybrid MPPT algorithm for the LLC resonant converter in PV power applications. IEEE Transactions on Power Electronics. 2021;37(1):944-54. 10.1109/TPEL.2021.309467
5. 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. 10.5829/ije.2023.36.06c.08
6. Zeb K, Khan I, Uddin W, Khan MA, Sathishkumar P, Busarello TDC, et al. A review on recent advances and future trends of transformerless inverter structures for single-phase grid-connected photovoltaic systems. Energies. 2018;11(8):1968. 10.3390/en11081968
7. Giacomini JC, Michels L, Cavalcanti MC, Rech C. Modified discontinuous PWM strategy for three-phase grid-connected PV inverters with hybrid active–passive damping scheme. IEEE Transactions on Power Electronics. 2019;35(8):8063-73. 10.1109/TPEL.2019.2961648
8. Bharath GV, Hota A, Agarwal V. A new family of 1-φ five-level transformerless inverters for solar PV applications. IEEE Transactions on Industry Applications. 2019;56(1):561-9. 10.1109/TIA.2019.2943125
9. Guo X, Yang Y, Wang B, Blaabjerg F. Leakage current reduction of three-phase Z-source three-level four-leg inverter for transformerless PV system. IEEE Transactions on Power Electronics. 2018;34(7):6299-308. 10.1109/TPEL.2018.2873223
10. Hosseinpour M, Rasekh N. A single-phase grid-tied PV based trans-z-source inverter utilizing LCL filter and grid side current active damping. Journal of Energy Management and Technology. 2019;3(3):67-77. 10.22109/jemt.2019.169380.1150
11. Aalizadeh F, Hosseinpour M, Dejamkhooy A, Shayeghi H. Two-stage control for small-signal modeling and power conditioning of grid- connected quasi-Z-Source inverter with LCL filter for photovoltaic generation. Journal of Operation and Automation in Power Engineering. 2021;9(3):242-55. 10.22098/joape.2021.7674.1546
12. Wu W, Liu Y, He Y, Chung HS-H, Liserre M, Blaabjerg F. Damping methods for resonances caused by LCL-filter-based current-controlled grid-tied power inverters: An overview. IEEE Transactions on Industrial Electronics. 2017;64(9):7402-13. 10.1109/TIE.2017.2714143
13. Rasekh N, Hosseinpour M. LCL filter design and robust converter side current feedback control for grid-connected Proton Exchange Membrane Fuel Cell system. International Journal of Hydrogen Energy. 2020;45(23):13055-67. 10.1016/j.ijhydene.2020.02.227
14. Giacomini JC, Michels L, Pinheiro H, Rech C. Design methodology of a passive damped modified LCL filter for leakage current reduction in grid-connected transformerless three-phase PV inverters. IET Renewable Power Generation. 2017;11(14):1769-77. 10.1049/iet-rpg.2017.0256
15. Dannehl J, Fuchs FW, Thøgersen PB. PI state space current control of grid-connected PWM converters with LCL filters. IEEE transactions on power electronics. 2010;25(9):2320-30. 10.1109/TPEL.2010.2047408
16. Liu J, Zhou L, Molinas M. Damping region extension for digitally controlled LCL-type grid-connected inverter with capacitor-current feedback. IET Power Electronics. 2018;11(12):1974-82. 10.1049/iet-pel.2018.0039
17. Rasekh N, Rahimian MM, Hosseinpour M, Dejamkhooy A, Akbarimajd A, editors. A step by step design procedure of PR controller and capacitor current feedback active damping for a LCL-type grid-tied T-type inverter. 2019 10th International Power Electronics, Drive Systems and Technologies Conference (PEDSTC); 2019: IEEE. 10.1109/PEDSTC.2019.8697853
18. Hosseinpour M, Asad M, Rasekh N. A step-by-step design procedure of a robust control design for grid-connected inverter by LCL filter in a weak and harmonically distorted grid. Iranian Journal of Science and Technology, Transactions of Electrical Engineering. 2021;45:843-59. 10.1007/s40998-021-00414-z
19. Dragičević T, Zheng C, Rodriguez J, Blaabjerg F. Robust quasi-predictive control of $ LCL $-filtered grid converters. IEEE Transactions on Power Electronics. 2019;35(2):1934-46. 10.1109/TPEL.2019.2916604
20. Cai Y, He Y, Zhou H, Liu J. Active-damping disturbance-rejection control strategy of LCL grid-connected inverter based on inverter-side-current feedback. IEEE Journal of Emerging and Selected Topics in Power Electronics. 2020;9(6):7183-98. 10.1109/JESTPE.2020.3017678
21. Guan Y, Wang Y, Xie Y, Liang Y, Lin A, Wang X. The dual-current control strategy of grid-connected inverter with LCL filter. IEEE Transactions on Power Electronics. 2018;34(6):5940-52. 10.1109/TPEL.2018.2869625
22. Zhou X, Zhou L, Chen Y, Shuai Z, Guerrero JM, Luo A, et al. Robust grid-current-feedback resonance suppression method for LCL-type grid-connected inverter connected to weak grid. IEEE Journal of Emerging and Selected Topics in Power Electronics. 2018;6(4):2126-37. 10.1109/JESTPE.2018.2805823
23. He Y, Wang X, Ruan X, Pan D, Qin K. Hybrid active damping combining capacitor current feedback and point of common coupling voltage feedforward for LCL-type grid-connected inverter. IEEE Transactions on Power Electronics. 2020;36(2):2373-83. 10.1109/TPEL.2020.3008160
24. Faiz MT, Khan MM, Jianming X, Ali M, Habib S, Hashmi K, et al. Capacitor voltage damping based on parallel feedforward compensation method for LCL-filter grid-connected inverter. IEEE Transactions on Industry Applications. 2019;56(1):837-49. 10.1109/TIA.2019.2951115
25. Li X, Fang J, Tang Y, Wu X, Geng Y. Capacitor-voltage feedforward with full delay compensation to improve weak grids adaptability of LCL-filtered grid-connected converters for distributed generation systems. IEEE Transactions on Power Electronics. 2017;33(1):749-64. 10.1109/TPEL.2017.2665483
26. Zou C, Liu B, Duan S, Li R, editors. A feedfoward scheme to improve system stability in grid-connected inverter with LCL filter. 2013 IEEE Energy Conversion Congress and Exposition; 2013: IEEE. 10.1109/ECCE.2013.6647299
27. Liu B, Wei Q, Zou C, Duan S. Stability analysis of LCL-type grid-connected inverter under single-loop inverter-side current control with capacitor voltage feedforward. IEEE Transactions on Industrial Informatics. 2017;14(2):691-702. 10.1109/TII.2017.2766890
28. Figueredo RS, Matakas L. Integrated common and differential mode filter with capacitor-voltage feedforward active damping for single-phase transformerless PV inverters. IEEE Transactions

on Power Electronics. 2019;35(7):7058-72. 10.1109/TPEL.2019.2956184

29. Rodriguez-Diaz E, Freijedo FD, Vasquez JC, Guerrero JM. Analysis and comparison of notch filter and capacitor voltage feedforward active damping techniques for LCL grid-connected converters. IEEE Transactions on Power Electronics. 2018;34(4):3958-72. 10.1109/TPEL.2018.2856634

30. Said-Romdhane MB, Naouar MW, Slama Belkhodja I, Monmasson E. An improved LCL filter design in order to ensure stability without damping and despite large grid impedance variations. Energies. 2017;10(3):336. 10.3390/en10030336

Persian Abstract

چکیده

اینورترهای متصل به شبکه به منظور ارتباط موثر منابع انرژی تجدیدپذیر و کاربرد در سیستم تولید پراکنده به عنوان عناصر حیاتی در نظر گرفته می‌شوند. هارمونیک‌های جریان تولید شده ناشی از ریپل در لینک DC و کلیدزنی فرکانس بالا از جمله معایب اینورترهای متصل به شبکه هستند که با استفاده از فیلترهای LCL کاهش می‌یابند. با این وجود، تشدید ذاتی در فیلتر LCL منجر به ناپایداری سیستم انتقال توان می‌شود. در نتیجه، میراسازی مناسب برای حذف تشدید در فیلتر LCL حائز اهمیت می‌باشد. هدف این مقاله، بهبود کیفیت توان تزریقی آرایه فتوولتائیک متصل به شبکه مبتنی بر فیلتر LCL است. بدین منظور، پایداری سیستم اینورتر متصل به شبکه با به کارگیری روش میراسازی فعال بررسی شده و ردیابی نقطه حداکثر توان (MPPT) برای آرایه PV انجام شده است. روش میراسازی فعال پیشخور ولتاژ خازن با در نظر گرفتن تاخیر محاسباتی در این مقاله ارائه شده است. با ترکیب این روش و میراسازی فیدبک جریان سمت اینورتر، کنترل پیشنهادی مشخصات فرکانس پایین سیستم را مستقل از تغییرات امپدانس شبکه حفظ نموده و قابلیت رد هارمونیک بالا را بدون استفاده از جبرانسازهای اضافی فراهم می‌نماید. ولتاژ خازن نیز به عنوان جایگزینی از ولتاژ شبکه برای حلقه قفل فاز (PLL) اندازه‌گیری شده که منجر به کاهش تعداد حسگر شده است. در ضمن ردیابی نقطه حداکثر توان با استفاده از تکنیک هدایت افزایشی (IC) در مبدل بوست پیاده‌سازی شده است. علاوه بر این، روش محاسباتی ساده و مناسب برای طراحی پارامترهای فیلتر LCL ارائه شده و حساسیت سیستم تحلیل شده است. نهایتا، شبیه‌سازی در نرم افزار MATLAB پیاده‌سازی شده است که نتایج حاصله بیانگر عملکرد صحیح و دقیق سیستم کنترلی در تزریق حداکثر توان آرایه فتوولتائیک به شبکه و کیفیت بسیار مطلوب جریان تزریقی به شبکه است.

IJE TRANSACTIONS B: Applications Vol. 37 No. 02, (February 2024) 268-282

International Journal of Engineering

Journal Homepage: www.ije.ir

Seismic Fragility of FRC Columns using Incremental Dynamic Analysis and eXtended Finite Element Method

S. El Yassari*, A. EL Ghoulbzouri, S. El Janous

MODSGC Unit, National School of Applied Sciences Al Hoceima, University Abdelmalek Essaadi, Tangier, Morocco

PAPER INFO

Paper history:
Received 24 July 2023
Received in revised form 26 September 2023
Accepted 27 September2023

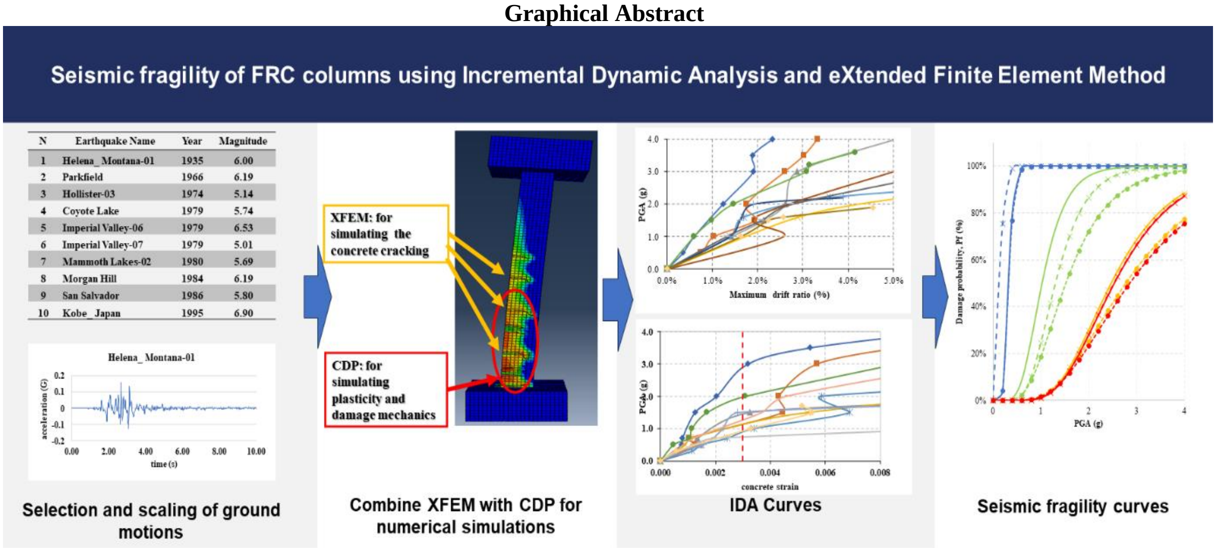
Keywords:
Fiber Reinforced Concrete
Incremental Dynamic Analysis
Pushover Analysis
Concrete Damage Plasticity
XFEM
Extended Finite Element Method

ABSTRACT

Ensuring seismic resilience in earthquake-prone regions is imperative for structural safety. Fiber-Reinforced Concrete (FRC) columns hold promise for enhancing structural performance under seismic conditions. This study seeks to comprehensively evaluate their seismic behavior. The primary objective of this research is to assess and compare the seismic performance of various FRC column types, including polypropylene fibers (PFRC), steel fibers (SFRC), and hybrid combinations (HyFRC), in contrast to conventional reinforced concrete (RC) columns. To achieve this, the study employs eXtended Finite Element Method combined with Concrete Damage Plasticity (XFEM-CDP) in Abaqus to scrutinize static and dynamic responses. The nonlinear static pushover analysis unveiled a notable improvement in seismic resistance across all FRC types when compared to RC columns. Incremental dynamic analyses (IDA) are conducted using the selected suite of 10 near fault as-recorded ground motions to evaluate the inelastic seismic responses of different FRC bridge columns. XFEM-CDP simulations in Abaqus captured multiple aspects of FRC columns, such as concrete cracking, loss of stiffness and plastic behavior. Seismic fragility analysis of these FRC columns is conducted considering four damage states: a) longitudinal steel yielding, b) core concrete crushing, c) steel bar buckling, and d) longitudinal steel bar fracture. The results indicated that HyFRC columns exhibit the lowest damage vulnerability compared to PFRC and SFRC variants.

doi*: 10.5829/ije.2024.37.02b.05*

Graphical Abstract



Seismic fragility of FRC columns using Incremental Dynamic Analysis and eXtended Finite Element Method

N	Earthquake Name	Year	Magnitude
1	Helena_ Montana-01	1935	6.00
2	Parkfield	1966	6.19
3	Hollister-03	1974	5.14
4	Coyote Lake	1979	5.74
5	Imperial Valley-06	1979	6.53
6	Imperial Valley-07	1979	5.01
7	Mammoth Lakes-02	1980	5.69
8	Morgan Hill	1984	6.19
9	San Salvador	1986	5.80
10	Kobe_ Japan	1995	6.90

Selection and scaling of ground motions

XFEM: for simulating the concrete cracking

CDP: for simulating plasticity and damage mechanics

Combine XFEM with CDP for numerical simulations

IDA Curves

Seismic fragility curves

---

[1] Corresponding author email: *soufiane.elyassari1@etu.uae.ac.ma* (S. El Yassari)

Please cite this article as: El Yassari S, EL Ghoulbzouri A, El Janous S. Seismic Fragility of FRC Columns using Incremental Dynamic Analysis and eXtended Finite Element Method. International Journal of Engineering, Transactions B: Applications. 2024;37(02):268-82.

1. INTRODUCTION

Reinforced Concrete (RC) bridges are extensively built across the globe, with a significant portion of them situated in regions prone to seismic activity. In compliance with contemporary seismic regulations, these structures are mandated to maintain adequate capacity to endure a catastrophic earthquake. Nonetheless, recent substantial seismic events, such as the 2011 Japan earthquake (Mw 9.0) and the 2023 Turkey earthquake (Mw 7.8), resulted in extensive damage. This damage can be attributed, in part, to the inherent limitations of plain concrete, which includes a lack of flexural strength, brittleness, low toughness, and limited energy absorption capabilities.

Traditionally, increasing transverse stirrups in bridge columns enhances concrete properties. However, this can lead to construction problems. An alternative solution gaining traction is fiber-reinforced concrete (FRC). FRC integrates various fiber types like steel and polypropylene, improving tensile strength, toughness, and flexural strength. It also enhances ductility and energy absorption during seismic events.

Various studies investigate FRC in bridge columns, including steel fiber (SFRC), polypropylene fiber (PFRC), and hybrid fiber (HyFRC) types. SFRC enhances ductility and loading capacity, with an optimal volume fraction of 1.5% for shear strength (1). PFRC boosts ductility and energy absorption. HyFRC shows superior dynamic capacity, especially at higher axial compression ratios.

Existing research on FRC columns lacks a holistic evaluation under diverse seismic scenarios and materials. Additionally, the comparative advantages of different fiber types for seismic resilience have not been fully explored. Seismic analyses have conventionally relied on linear and simplified static or dynamic methods, which may not fully represent the intricate interactions between seismic forces and structural behavior in reality. Probabilistic approaches, such as the development of fragility curves, have emerged as a promising alternative for seismic assessments (2-5).

Over the past two decades, the field of structural seismic assessment has rapidly advanced with the probabilistic approach. Fragility curves, which estimate the interdependence between structural demands and capacity, have become standard in seismic assessments (3). Various methodologies have been proposed for generating structural seismic fragility curves, encompassing empirical and analytical approaches (3-8). Analytical methods, including elastic-spectral (3-5), non-linear static (7-10), and non-linear dynamic analyses (11-15), have emerged as diverse and numerous fragility generation methodologies.

The performance-based earthquake engineering (PBEE) procedure necessitates evaluating the interplay between structural capacity and seismic demands. To address this, the Incremental Dynamic Analysis (IDA) was introduced by Luco and Cornell (16), further developed by Vamvatsikos and Cornell (17), enabling the prediction of a structure's response ranging from elastic behavior to inelastic behavior and ultimately collapse.

Non-linear methods, such as Nonlinear Static Pushover Analysis (NSPA) and IDA, are widely employed due to their relative accuracy in structural design and seismic analysis. Pushover analysis, a non-linear static method, offers a quick evaluation of a structure's seismic performance (18, 19). In contrast, IDA, incorporating dynamic effects into nonlinearity, provides more accurate results than non-linear static analysis. Although non-linear dynamic methods require more computational time and effort, advancements in computing power have increased their popularity in recent years.

Traditional methods of seismic analysis have often proven inadequate in capturing the complex nonlinear behavior of structures subjected to strong earthquakes. To address this limitation and improve the accuracy of seismic assessments, researchers have been pioneering an innovative approach by combining Computational Damage Plasticity (CDP) and the eXtended Finite Element Method (XFEM) within the Abaqus software framework. This synergistic integration allows for a comprehensive evaluation of structural response under seismic loading conditions, incorporating both NSPA and IDA.

This paper aims to assess of the seismic performance of FRC columns. 3D nonlinear finite element models have been constructed within the Abaqus software to simulate the seismic behavior of these FRC columns.

These models are first calibrated using available experimental data from the literature. The seismic

TABLE 1. Summary of the material properties

Material	Diameter (mm)	Length (mm)	Density (g/cm^3)	Tensile /compressive strength (MPa)	Elastic modulus (GPa)
Concrete	-	-	2.5	40	30
Longitudinal steel	20	-	7.85	400	200
Hoop steel	12	-	7.85	335	200
Steel fibers	0.55	35	7.85	1143	200
Polypropylene fibers	0.048	19	1.2	400	43

resilience of columns employing various FRC materials is evaluated by considering four flexural damage states: longitudinal reinforcement yielding, concrete crushing, buckling of reinforcement, and longitudinal reinforcement fracture. These damage states are derived through NSPA.

Furthermore, IDA is conducted to examine the seismic demands placed on FRC columns subjected to intense ground motions. In the IDA approach, a selection of ten near-fault ground motions is employed as seismic inputs for the finite element models. These ground motions are scaled to different intensity levels. The outcomes of IDA are presented in the form of IDA curves, illustrating the relationship between the engineering demand parameter (EDP) of the structures and the intensity measure (IM) of the ground motions. For this study, peak ground acceleration (PGA) is adopted as the IM. Various EDPs of the bridge columns, such as maximum drift, longitudinal steel strain, and concrete strain demands, were identified and documented throughout the IDA analysis. finally, the seismic fragility curves were constructed and compared to explore the impact of different FRC materials on the seismic performance of the columns.

2. MODELING OF DIFFERENT FRC COLUMNS

2. 1. Details of Columns Figure 1 provides the dimensions of the studied column, which consists of 20 longitudinal reinforcement bars with a diameter of 20 mm, resulting in a steel-to-concrete ratio of 2%. The column has a height of 6 meters and a square cross-section measuring 0.6 m by 0.6 m. The pile cap dimensions are 1.8 m by 1.8 m by 0.8 m, while the foundation measures 1.8 m by 1.8 m by 0.8 m. The square transverse stirrups have a diameter of 12 mm and are spaced at intervals of 150 mm.

This study focuses on four different types of columns: RC (conventional reinforced concrete), SFRC (steel fiber-reinforced concrete), PFRC (polypropylene fiber-reinforced concrete), and HyFRC (a hybrid combination of steel and polypropylene fibers). The material characteristics for each type are presented in Table 1. The parameters for the steel fibers were adopted from literature (20), while those for the polypropylene fibers were derived from literature (21). The steel fibers have a tensile strength of 1143 MPa, an elastic modulus of 200 GPa, and a density of 7.8 g/cm³.

To ensure the workability of the concrete mixture is not compromised, the recommended optimal volume fraction of steel fibers in the SFRC columns is 1.5% (117 kg/m³). The PFRC columns contain a volume fraction of 0.15% (1.37 kg/m³) of polypropylene fibers. Previous research by Huang et al. (21), Yin et al. (22) suggested that the optimal fiber combination for enhancing concrete strength in the HyFRC is 1.5% steel fibers and 0.15% polypropylene fibers. Table 2 provides the main parameters for the various FRC columns.

2. 2. Finite Element Modeling Finite element models of columns made of standard RC and FRC materials are created using the commercial software Abaqus. Truss elements are used to simulate the rebars, and solid three-dimensional eight-node linear brick with full integration (C3D8) simulates the concrete elements. The dynamic implicit algorithm Newton method with transient fidelity is used to implement the time-history analysis. The concrete damaged plasticity model (CDP) (23, 24) has been used to describe the stiffness degradation of concrete during cyclic loading, which can be defined using two factors, dt and dc (Figure 2). The compressive and tensile damage factors, dc and dt are estimated based on the corresponding inelastic strains using Equations 17-22 (25).

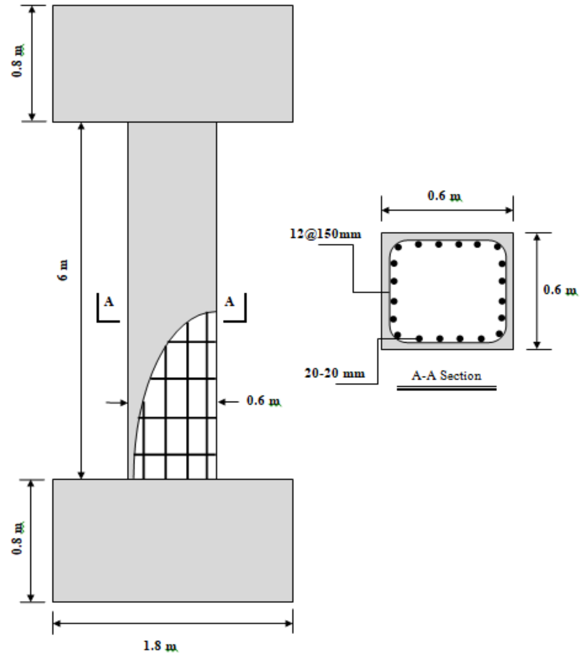
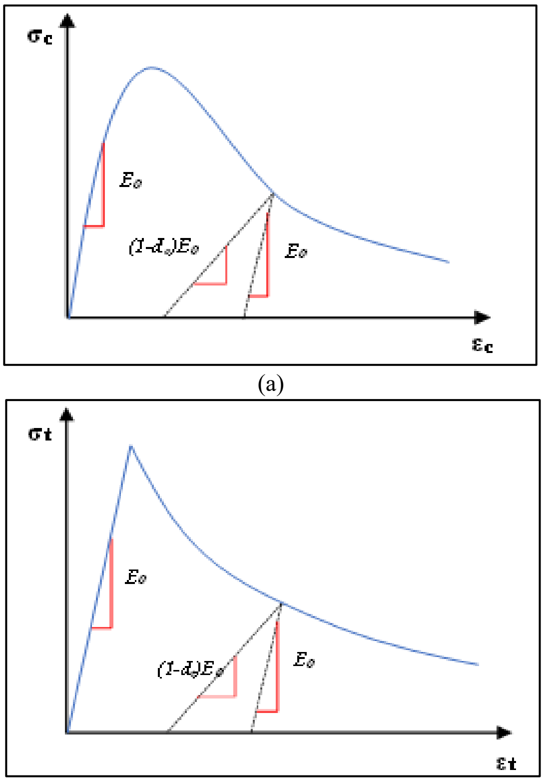


Figure 1. Column dimensions and reinforcement details

TABLE 2. Main parameters for different bridge columns

Material	Compressive strength (MPa)	Steel fiber content	Polypropylene fiber content
RC	40	-	-
SFRC	45	1.50%	-
PFRC	43	-	0.15%
HyFRC	48	1.50%	0.15%



(a)

(b)

Figure 2. Stress-strain relationships of CDP in ABAQUS, (a) compressive curve, (b) tensile curve

The ABAQUS software simulates crack propagation using the eXtended Finite Element Method (XFEM), while material nonlinearities are accounted for using CDP criteria. XFEM is used in the simulation to solve the discontinuity problems.

In practice, discontinuities can manifest as flaws or cracks. There are various kinds of discontinuity:

- A strong discontinuity represents the cracks.
- A weak discontinuity represents the interface between two different materials.

It is challenging to analyze the discontinuity in a concrete model using the traditional finite element method. Without the need for remeshing, XFEM can simulate the initiation and propagation of a discrete crack along any solution path. The combination of XFEM and CDP in this paper is presented in Figure 3.

2. 3. Material Property

2. 3. 1. Compressive Behavior

The compressive behavior model in this study is based on Abadel et al.'s (26) work:

$$\sigma_c = \left(\frac{\beta\left(\frac{\varepsilon_c}{\varepsilon_0}\right)}{\beta - 1 + \left(\frac{\varepsilon_c}{\varepsilon_0}\right)^{\beta}} \right) f_{cu} \qquad (1)$$

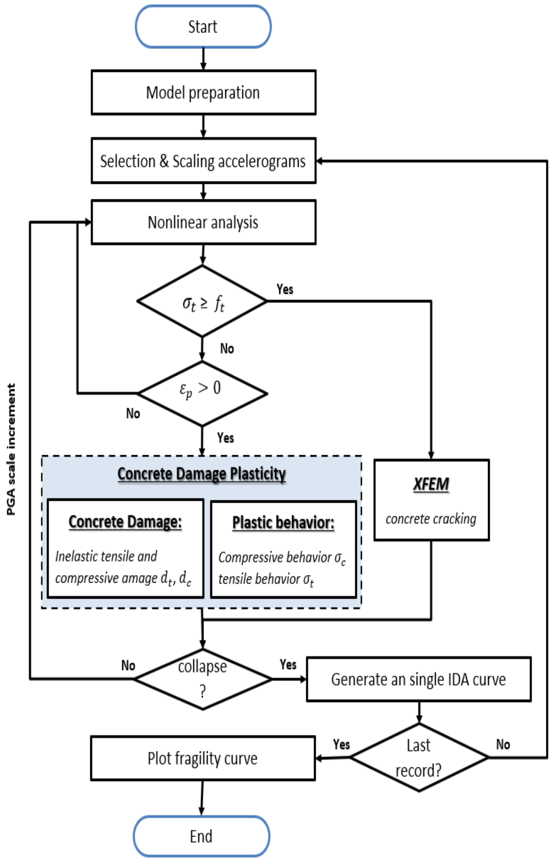


Figure 3. Flowchart of combined XFEM/CDP method

where f_{cu}denotes the ultimate compressive stress, ε_c and ε_0 represents the compressive strain and the strain at peak stress of plain concrete (= 0,002), respectivly.

The parameters can be calculated using the following equations:

$$\beta = 1 + 5e^{-1,376RI_v} \qquad (2)$$

$$\beta_0 = 0,108f_c - 0,966 \qquad (3)$$

$$f_{cfrc} = f_{c0} + 5,222RI_v \qquad (4)$$

$$\varepsilon_{frc} = \varepsilon_{c0} + 0,0004RI_v \qquad (5)$$

$$RI_v = \sum RI_{vi} \qquad (6)$$

$$RI_{vi} = V_i \frac{l_i}{d_i} \frac{E_i}{E_s} \qquad (7)$$

For the plain concrete, the value of parameter $\beta_0$can be calculated with the help of Equation 3. RI_v, V_i,l_i, and d_i denote thereinforcing index, fibers' volume fraction, length, and diameter (or their corresponding diameter in non-circular sections), E_i and E_s are the fiber's and steel material's modulus of elasticity, respectively.

2. 3. 2. Tensile Behavior The tensile stress-crack opening displacement relationship adopted for plain concrete in this paper has been proposed by Hordijk (27). whereas for HFRC, the constitutive relationship was developed by Almusallam et al. (28) which is based on the reinforcing index, RIv:

$$\frac{\sigma_t}{f_t}=\left(1+\left(c_1\frac{w_t}{w_{cr}}\right)^3\right)e^{\left(-c_1\frac{w_t}{w_{cr}}\right)}-\frac{w_t}{w_{cr}}(1+{c_1}^3)e^{(-c_1)} \tag{8}$$

$$f_t=1{,}4\left(\frac{f_{cu}-8}{10}\right)^{2/3} \tag{9}$$

$$G_F=\left(0{,}0469{d_a}^2-0{,}5d_a+26\right)\left(\frac{f_{cu}}{10}\right)^{0{,}7} \tag{10}$$

$$w_{cr}=5{,}14\frac{G_F}{f_t} \tag{11}$$

In these equations, the f_t is the tensile strengths of plain concrete, w_t, w_{cr} denotes the crack opening displacement and crack displacement at the complete loss of tensile stress, respectivly. d_a and l_{eq} represent the maximum aggregate size of the concrete (20 mm) and the mesh element length, respectivly.

Concerning the tensile behavior of FRC, $\sigma_t(w)$, Rousakis et al. (29) proposed an analytical model to describe the tensile softening behavior based on the reinforcing index, RIv. This model was obtained through inverse analysis and provides a good agreement with experimental results.

$$\sigma_t(w_t)=a_1f_{tfrc}e^{(-a_2w_t)} \tag{12}$$

$$a_1=0{,}75 \tag{13}$$

$$a_2=10e^{-4{,}3RI_v} \tag{14}$$

$$f_{tfrc}=(f_t-1)+e^{1{,}23RI_v} \tag{15}$$

where f_{tfrc} denotes the tensile strength of FRC.

It was assumed that each element has a single crack. thus, the relationship between strain and the crack opening displacement is defined by Equation 14.

$$\varepsilon_t=\varepsilon_{tm}+w_t/l_{eq} \tag{16}$$

ε_{tm} is the tensile strain corresponding to the tensile strengths.

2. 3. 3. Concrete Damage Plasticity Model The compressive and tensile damage factors, dc and dt are given by Equations 15-20.

$$d_t=\frac{1}{e^{-\frac{1}{m_t}}-1}\left(e^{-\frac{\varepsilon_{t,nom}^{ck}}{m_t}}-1\right) \tag{17}$$

$$d_c=\frac{1}{e^{-\frac{1}{m_c}}-1}\left(e^{-\frac{\varepsilon_{c,nom}^{in}}{m_c}}-1\right) \tag{18}$$

$$\varepsilon_{t,nom}^{ck}=\frac{\varepsilon_t^{ck}}{\varepsilon_{tu}^{ck}} \tag{19}$$

$$\varepsilon_{c,nom}^{in}=\frac{\varepsilon_c^{in}}{\varepsilon_{cu}^{in}} \tag{20}$$

m_c,m_t, are the parameters that control damage evolution speed. $\varepsilon_{t,nom}^{ck}$, $\varepsilon_{c,nom}^{in}$ denote the normalized compressive and tensile inelastic strains.ε_{tu}^{ck},ε_{cu}^{in}, are the corresponding ultimate strains.

Typically, for plain concrete, Chi et al. (25) suggested the use of $m_t=0{,}05$, $m_c=0{,}1$, $\varepsilon_{cu}^{in}=0{,}033$, $\varepsilon_{tu}^{ck}=0{,}0033$.

For FRC, parameters m_c^{hf} and m_t^{hf} are modified according to the composite material theory as follows:

$$m_c^{hf}=m_c\left(1+a_{m1}\lambda_{sf}+b_{m1}\lambda_{pf}\right) \tag{21}$$

$$m_t^{hf}=m_t\left(1+a_{m2}\lambda_{sf}+b_{m2}\lambda_{pf}\right) \tag{22}$$

where λ_{sf} and λ_{pf} are the characteristic parameters of steel and polypropylene fibers, respectively.

The values am1 = 0.452, bm1 = 0.054, am2 = 0.628, and bm2 = 0.156 were recommended by Chi et al. (25).

The key parameters, K_c^{hf}, $\sigma_{b0}^{hf}/\sigma_{c0}^{hf}$ and the dilation angle ψ^{hf} are defined based on the following equations:

$$K_c^{hf}=K_c.\frac{k_t}{k_c} \tag{23}$$

$$k_t=1+0{,}080\lambda_{sf}+0{,}132\lambda_{pf} \tag{24}$$

$$k_c=1+0{,}056\lambda_{sf} \tag{25}$$

$$\lambda_{sf}=V_{sf}\frac{l_{sf}}{d_{sf}} \tag{26}$$

$$\lambda_{pf}=V_{pf}\frac{l_{pf}}{d_{pf}} \tag{27}$$

$$\frac{\sigma_{b0}^{hf}}{\sigma_{c0}^{hf}}=\frac{{k_t}^2}{0{,}132k_c}\left[\left(0{,}728-\frac{0{,}749}{k_t}\right)+\sqrt{\left(0{,}728-\frac{0{,}749}{k_t}\right)^2+\frac{0{,}03}{{k_t}^2}}\right] \tag{28}$$

$$\psi^{hf}=\psi_0(1-a_\psi\lambda_{sf}-b_\psi\lambda_{pf}) \tag{29}$$

$$\psi^{hf}=\psi_0(1-a_\psi\lambda_{sf}-b_\psi\lambda_{pf}) \tag{30}$$

$$\psi_0=36+(\sigma_{c0}/\sigma_{cm0}) \tag{31}$$

In this equation, σ_{c0} is a parameter that ensures the equivalence of units, and σ_{cm0} is recommended to be 10 MPa according to (29). The coefficients a_ψ and b_ψ, which are obtained from (25), have values of 0.861 and 0.097, respectively. These coefficients are used to calculate the dilation angle for plain concrete in the proposed model.

2. 3. 4. Concrete Crack Evolution Law In standard FEM, cracks must follow element edges or a predefined path, such as the cohesive zone model. In contrast, the XFEM does not need the crack path to be predefined, making it independent of the model mesh. In the XFEM, Crack initiation criteria must be specified. This study uses the ultimate tensile strength ft as the crack initiation criteria. On the other hand, crack evolution is determined based on the displacement parameter and Equations 8 and 12, as shown in Figure 4. The XFEM can be implemented in Abaqus Implicit to simulate the concrete cracking process using the parameters listed in Table 3.

2. 3. 5. Steel Material Model

The reinforcing steel bars in the concrete columns were modeled as truss elements (T3D2) with linear elastic behavior, followed by linear hardening until reaching the ultimate stress in tension (σu). The strain-softening and the fracture of the bars are simulated in Abaqus using the damage fracture option, where (wf) is the estimated displacement at failure based on calibration with test results (Figure 5). For simplification purposes, the bond between concrete and steel is assumed to be perfect. The debonding of steel is simulated as steel degradation. Table 4 shows the mechanical properties used in the reinforcing steel material model.

2. 3. 6. Model Validation with Experimental Results from the Literature

Quasi-static cyclic pushover analyses were performed using the Abaqus software to simulate the behavior of the columns. Finite element models of columns incorporating different RC and FRC materials were calibrated and validated using experimental data from previous studies (20, 21, 30). The main parameters for the various FRC columns are summarized in Table 5.

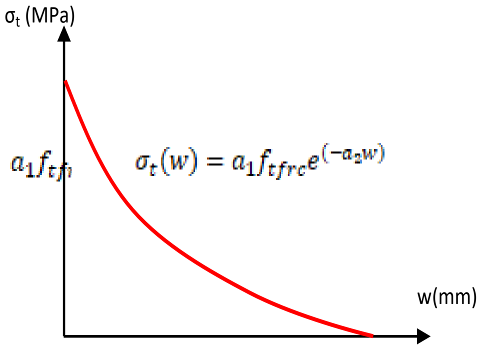


Figure 4. Tensile strength and Crack evolution relationship for FRC

TABLE 3. XFEM parameters

	RC	FRC
Crack initiation	ft	a1*ftFRC
Crack evolution law	Exponential	Exponential
evolution law parameter	c1/wcr	a2
Displacement at failure	wcr	2/a2

Constitutive models for the steel bars and concrete matrix components were separately established for numerical modeling. Truss elements (T3D2) were used to model the stirrups, while a solid three-dimensional eight-node linear brick with full integration (C3D8) was assigned to represent the concrete matrix. It is worth noting that full integration elements (C3D8) were utilized as they converge faster with the eXtended Finite Element Method (XFEM) compared to reduced integration elements (C3D8R).

Figure 6 illustrates the hysteretic curves and corresponding skeleton curves of the ten different types of RC, SFRC, PFRC, and HyFRC columns. The solid black line represents the experimental results, while the dashed red line represents the numerical simulation results. The figure demonstrates the cyclic deterioration of these columns' performance, including the gradual decrease in unloading stiffness, the degraded strength resulting from cyclic loading, and the pinching effect caused by concrete cracking.

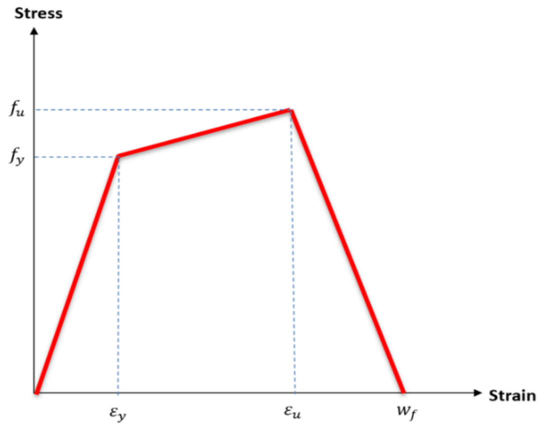


Figure 5. Reinforcing steel model for steel bars

TABLE 4. Steel parameters for different models

Model	d (mm)	fy (MPa)	fu (MPa)	Es (MPa)	εh	εu
(a). Zhang2013	10	335	500	2.0*105	0.001675	0.06
(b). Zhang2017	14	335	500	2.0*105	0.001675	0.06
(c). Huang2015	14	553.9	670.3	2.0*105	0.001675	0.06
(d). Liang2016	16	440	609	1.95*105	0.001675	0.08

TABLE 5. Values for CPDM for (a, b, c, d) normal RC, (e, f, g) SFRC, (h) PFRC and (i) HYFRC

Model	Vsf (%)	Vpf (%)	ψ (°)	$\sigma_{b0}^{hf}/\sigma_{c0}^{hf}$	K	Mesh size (mm)
(a). RC-Zhang2017	0.0	0.0	38.70	1.162	0.666	50
(b). RC-Liang2016	0.0	0.0	38.70	1.162	0.666	60
(c). SFRC-Zhang2017	1.0	0.0	16.775	1.46	0.676	50
(d). PF-1-1-Huang2015	0.0	0.15	34.968	1.162	0.718	40
(e)-(f). HF-1-1-Huang2015	1.5	0.15	4.085	1.641	0.731	40

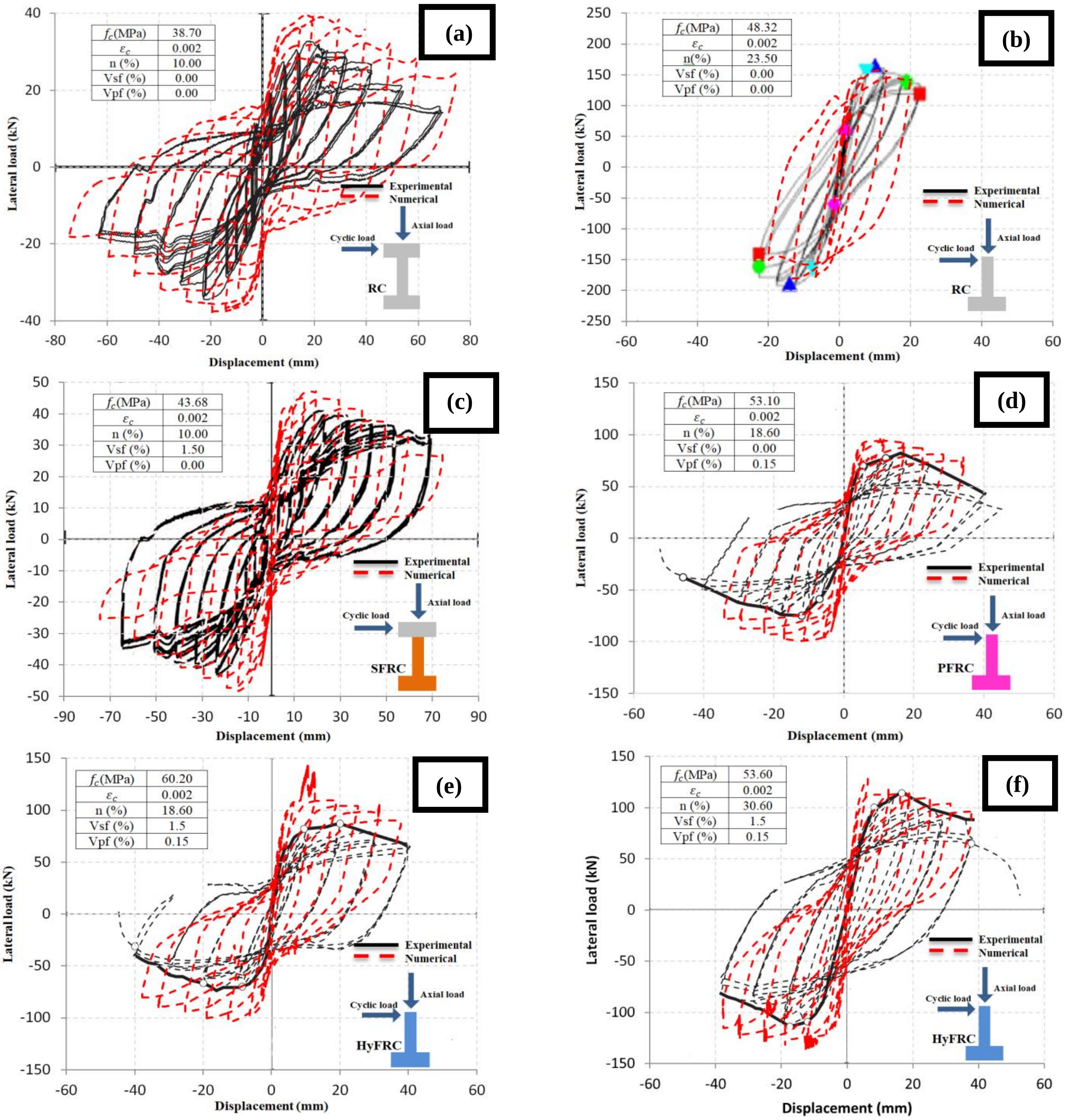


Figure 6. Comparison of the numerical hysteresis curve of different columns with the experimental hysteresis response: a, b) RC (12, 21), c) SFRC (12), d) PFRC (13) and e, f) HyFRC (13)

The comparison results are shown in Figure 6 indicate that the dynamic hysteretic model can provide reasonable estimates of the strength capacity of the various FRC columns. Furthermore, the proposed degradation parameters accurately depict the deterioration of strength and stiffness, as well as the pinching effect observed in the experimental data.

The analytical model introduced by Zhang and Dias-da-Costa (20), exhibited the most significant difference between the predicted load and the actual test results. This discrepancy can be primarily attributed to the utilization of a relatively large mesh size, as indicated in Table 5 (50 mm). Notably, larger mesh elements necessitate higher fracture energies, leading to more substantial inaccuracies in lateral load predictions. This highlights the significance of conducting a sensitivity analysis concerning mesh size, particularly since more pronounced errors tend to manifest at smaller displacements.

2. 4. Nonlinear Static Pushover Analysis

The finite element model included the weight load of the materials, with concrete density of 2500 kg/m^3 and steel density of 7850 kg/m^3. The superstructure load of 171 kN and lateral load were applied at the top of the pile cap. Consequently, the columns had an axial load ratio determined as 0.02 times the product of concrete strength (fc) and column section area (Ac). The results of the nonlinear static pushover analysis (NSPA) established the relationship between the shear base and lateral displacement, commonly known as pushover curves. The deformed shape and the tensile damage of different RC and FRC columns are presented in Figure 7.

Figure 8 illustrates the pushover curves of the different reinforced concrete (RC) and fiber-reinforced concrete (FRC) columns. As depicted, all FRC columns exhibited significantly increased capacity compared to the standard RC column. Among the FRC columns, the HyFRC column demonstrated the most substantial enhancement in seismic capacity, followed by the SFRC and PFRC columns. Moreover, the variations in pushover curves between the SFRC, PFRC, and HyFRC columns were minimal at small lateral displacements, primarily due to the similar elastic modulus of these FRC materials.

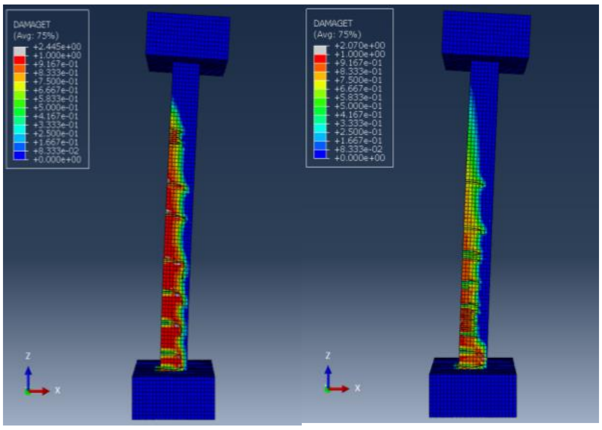
The differences between the SFRC and PFRC columns became less significant at larger lateral displacements, as the improvement in compressive strength offered by both types of fibers was comparable. Overall, all FRC materials exhibited a significant improvement in seismic capacity.

The structural damage classification according to HAZUS includes "Slight, Moderate, Extensive, and Collapse." This paper considers four flexural damage states: a) longitudinal steel yielding, b) core concrete crushing, c) steel bar buckling, and d) longitudinal steel bar fracture. The limit values for these damage states can be determined based on the strain limits of concrete and steel bars.

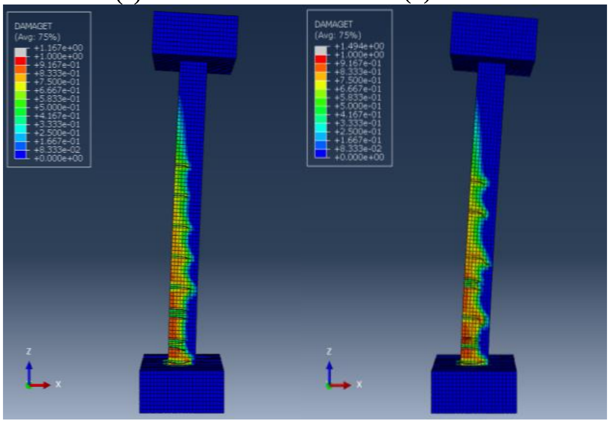
It is assumed that longitudinal steel bars yield when the steel strain reaches the ratio of yield stress to elastic modulus of the steel bars. According to Paulay and Priestley (31), core concrete crushing occurs when the concrete strain reaches 0.003, and sometimes higher values such as 0.006 to 0.008 are observed. The ultimate compression strain of 0.004 is recommended. ACI and AASHTO propose a design maximum strain of concrete crushing at 0.003. In this paper, a concrete crushing strain of 0.003 is adopted as the reference value. Furthermore, the buckling and fracture of steel bars can be predicted using the model projected by Berry and Eberhard (32) based on the tensile strain of longitudinal reinforcements.

Table 6 presents the lateral displacements and base shear obtained from the pushover curves at the four flexural damage states.

$$\varepsilon_{bb}^{calc} = 0.045 + 0.25\rho_{eff} \tag{32}$$



(a) RC (b) PFRC



(c) SFRC (d) HyFRC

Figure 7. Tensile damage results for different column types: a) RC, b) PFRC, c) SFRC and d) HyFRC

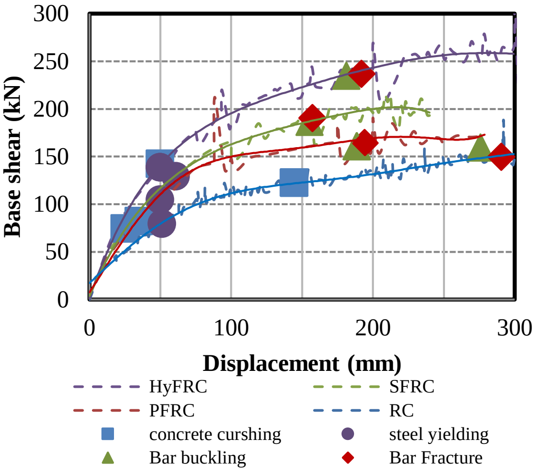


Figure 8. Pushover curves for different FRC columns

$$\varepsilon_{bf}^{calc} = 0.045 + 0.3\rho_{eff} \tag{33}$$

$$\rho_{eff} = \rho_s f_{ys}/f'_c \tag{34}$$

where ρ_{eff} is the effective confinement ratio, $\rho_s$is the spiral-reinforcement ratio, and $f_{ys}$is the yield stress of the spiral reinforcement.

Table 6 reveals that incorporating fibers into concrete enhances the shear base of columns. The slight damage state (longitudinal steel yielding) initiates at the same level of displacement in all columns for the steel yielding limit state. Normal RC columns have a base shear at steel yielding that is 79.25 kN, 31.8%, 62.7%, and 75.1% lower than SFRC, PFRC, and HyFRC columns, respectively. The base shear is enhanced in all columns for the FRC columns at extensive and collapse damage states, such as longitudinal bar buckling and fracture.

TABLE 6. Relative difference of base shear and displacement at different damage states

Material	**Steel yielding**		**Concrete crushing**		**Steel buckling**		**Steel fracture**	
	Displacement (mm)	**Base shear (kN)**	**Displacement (mm)**	**Base shear (kN)**	**Displacement (mm)**	**Base shear (kN)**	**Displacement (mm)**	**Base shear (kN)**
RC	51.31	79.25	144.46	122.5	275.95	159.2	290.67	149.5
PFRC	49.97	104.5	34.96	82.15	188.58	160.9	194.21	164
SFRC	60.87	129	25	74.23	155.44	186.2	157.4	190.4
HyFRC	49.74	138.8	49.95	142.1	181.22	234.6	191.84	236.8

3. INCREMENTAL DYNAMIC ANALYSIS

While NSPA can predict potential weak points in the structure under a lateral load, but it cannot be used to determine the dynamic effect of a given seismic excitation. On the other hand, the NTHA can predict the seismic demands with better accuracy. The IDA method is based on the NTHA for estimating significant inelastic responses of structures.

Damage measures such as drift ratio and displacement of the column's top are directly connected with the failure mechanism, which enables them to be considered as Engineering Demand Parameter (EDP) when generating IDA curves. On the other hand, the Peak Ground Acceleration (PGA) is selected as intensity measure (IM).

The IDA method is used in this paper to investigate the damage of columns and determine damage state limits under each ground motion. The steps required by the analysis (17):

1) Choose a set of records of seismic events that are representative of the hazard level;
2) Select a specific records and scale it using a scaling factor;
3) Perform a NTHA of the structure using this scaled record and identify the EDP and the IM of the ground motions;
4) Choose a set of records of seismic events that are representative of the hazard level;
5) Select a specific records and scale it using a scaling factor;
6) Perform a NTHA of the structure using this scaled record and identify the EDP and the IM of the ground motions;
7) Gradually increase IM using the scaling factor of this record until cover so that the simulations can cover the entire response range from the elastic behavior to total collapse;
8) Repeat Steps 2-4 for the same record;
9) Plot these points in order to generate the IDA curve

3. 1. Ground Motions

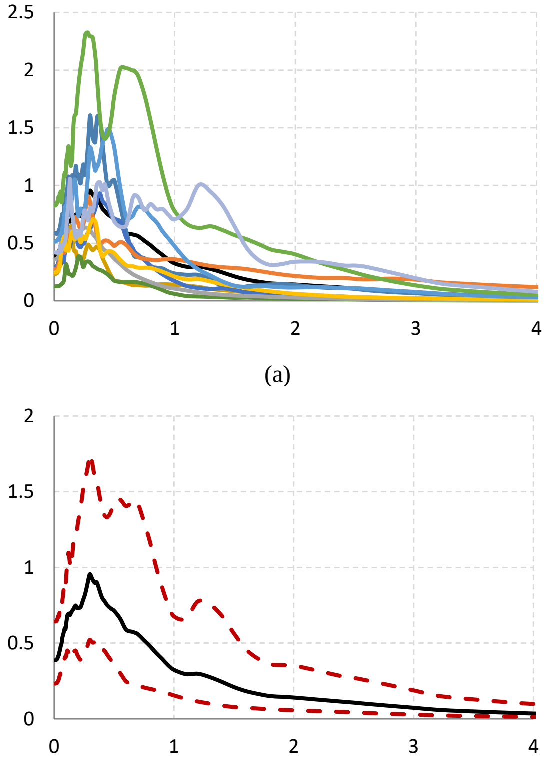
A set of 10 Earthquakes with magnitude ranging from 6.0 to 8.0 are selected from the PEER database and used to evaluate the seismic performance of columns, as summarized in Table 7. These ground motions have wave velocities Vs in the upper 30 m ranging from 200 m/s to 600 m/s. The records' closest site-to-fault distances range from 0.1 to 10 km. Figure 9 depicts the spectral accelerations of the ground motions, which demonstrated they can accurately represent medium to strong earthquakes.

3. 2. Maximum Drift Demand

A single Incremental Dynamic Analysis (IDA) curve is generated

TABLE 7. Selected ground motions used in IDA

N	Earthquake Name	Year	Station Name	Magnitude	Rrup (km)	Vs30 (m/sec)
1	Helena_ Montana-01	1935	Carroll College	6.00	2.86	593.35
2	Parkfield	1966	Cholame - Shandon Array #5	6.19	9.58	289.56
3	Hollister-03	1974	San Juan Bautista_ 24 Polk St	5.14	9.11	335.5
4	Coyote Lake	1979	Coyote Lake Dam - Southwest Abutment	5.74	6.13	561.43
5	Imperial Valley-06	1979	Brawley Airport	6.53	10.42	208.71
6	Imperial Valley-07	1979	El Centro Array #5	5.01	11.23	205.63
7	Mammoth Lakes-02	1980	Convict Creek	5.69	9.46	382.12
8	Morgan Hill	1984	Anderson Dam (Downstream)	6.19	3.26	488.77
9	San Salvador	1986	Geotech Investig Center	5.80	6.30	489.34
10	Kobe_ Japan	1995	Kobe University	6.90	0.92	1043



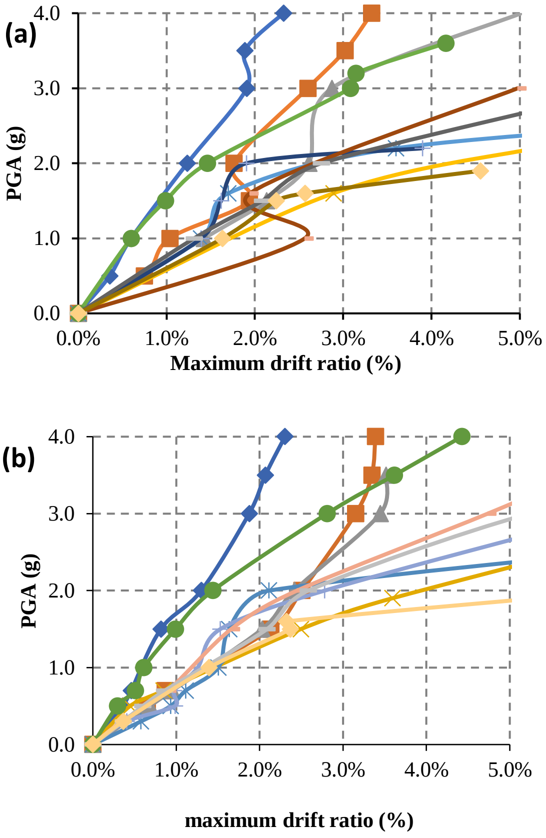
(a)

(b)

Figure 9. Selected as-recorded ground motions: a) spectral acceleration and b) percentiles of spectral

by measuring the maximum drift ratio of the column under incremental ground motion intensity levels. Figures 10-13 illustrates the IDA curves for SFRC, PFRC, and HyFRC columns, indicating both linear and nonlinear components that capture the elastic and inelastic behavior of structures subjected to varying Peak Ground Acceleration (PGA) intensities. The IDA curves effectively describe the softening behavior of the columns in the inelastic range until collapse occurs.

Due to the uncertainty of ground motions, the IDA curves exhibit diverse and scattered data, necessitating the use of statistical tools for evaluation. To provide more objective and reliable data, the IDA curves are analyzed using the 20th, 50th, and 80th fractile values. Figure 11



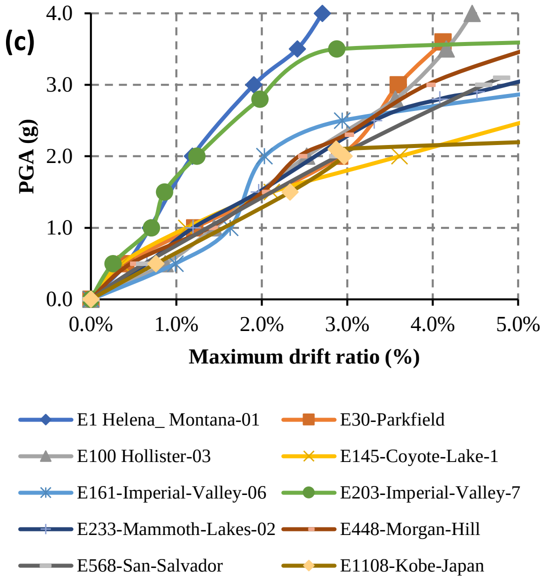


Figure 10. IDA curves for maximum drift (%) of different FRC columns: a) PFRC; b) SFRC and c) HyFRC

presents the median and percentiles of the maximum drift IDA curves for columns made of standard Reinforced Concrete (RC) and various Fiber-Reinforced Concrete (FRC) materials. It should be noted that for larger PGA values, the remaining number of IDA curves becomes small, making statistical calculation unrealistic. Therefore, the PGA range used for the fractile IDA curves is smaller than that of the original IDA curves.

From the results, it can be concluded that PFRC, SFRC, and HyFRC are all effective in improving seismic capacity and reducing seismic demands on columns. The improvement in seismic performance among the different FRC materials is negligible, with only a 5% difference observed. Previous studies (33, 34) have proposed acceptable drift values for serviceability and collapse damage states, with 1.9% indicating serviceability loss and 5.0% indicating global column collapse. Based on the obtained results, the maximum drift of all FRC columns remains within the collapse range, indicating an overall improvement in the seismic capacity of the columns.

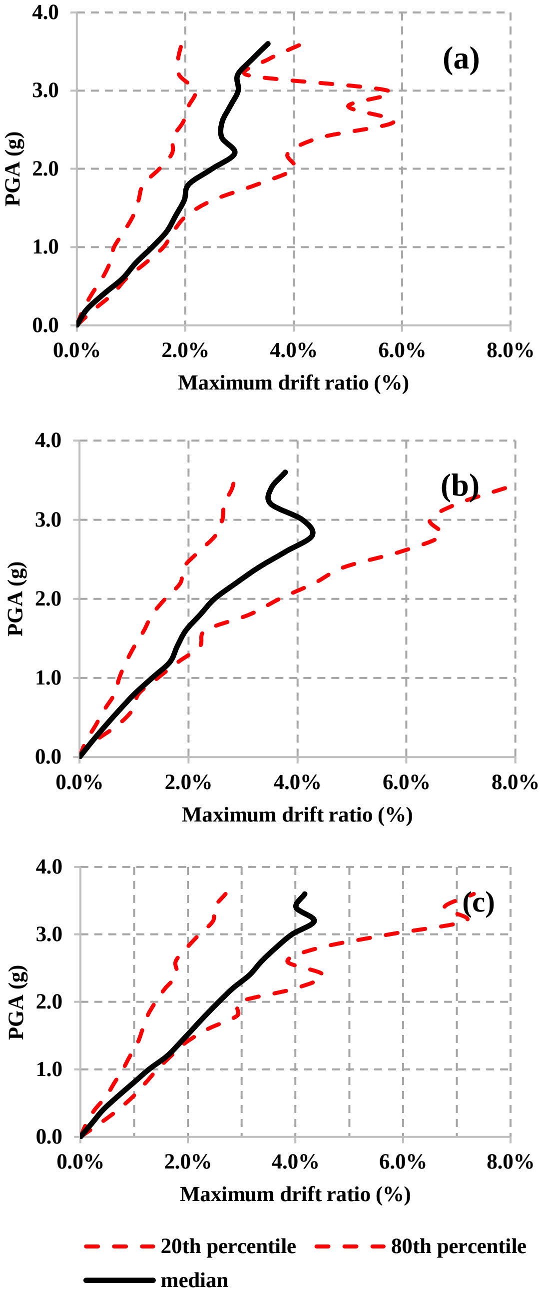


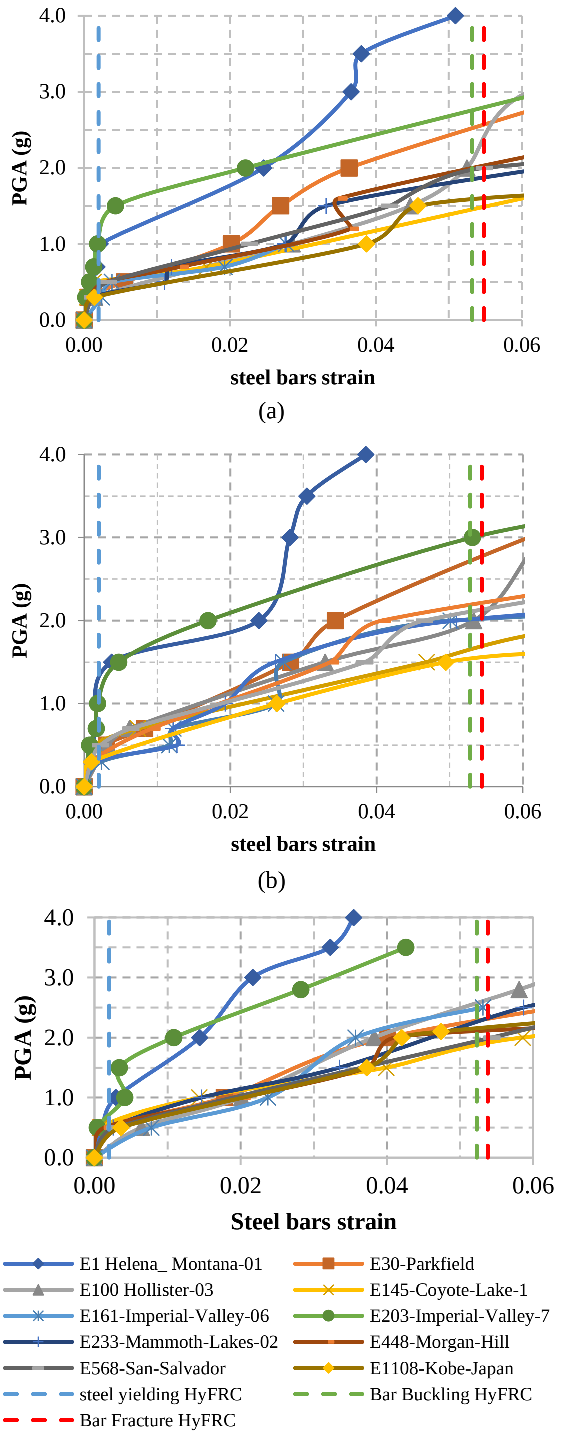
Figure 11. Median and percentiles of IDA curves for maximum drift (%) of different FRC columns: a) PFRC; b) SFRC and c) HyFRC

3. 3. Seismic Fragility

The seismic fragility analysis of the structures resulted in the development of fragility curves (FC), which represent the probability of surpassing a specific damage state at a given Intensity Measure (IM). These fragility curves were created using numerical simulations and employed two common methods: the Incremental Dynamic Analysis (IDA) and the cloud approach. The IDA involves incrementally scaling a selected set of earthquake records, while the cloud analysis involves unscaled records. By applying these methods, fragility functions were derived as: lognormal cumulative distribution functions (CDF) using Equation 34:

$$P_f(D \geq C|PGA = x) = \Phi(\frac{ln\,(x/c)}{\beta}) \quad (34)$$

In Equation 34, Pf represents the damage probability, PGA is the conditional value of the IM, and D and C denote the column demand and capacity, respectively. The function Φ(.) corresponds to the normal standard

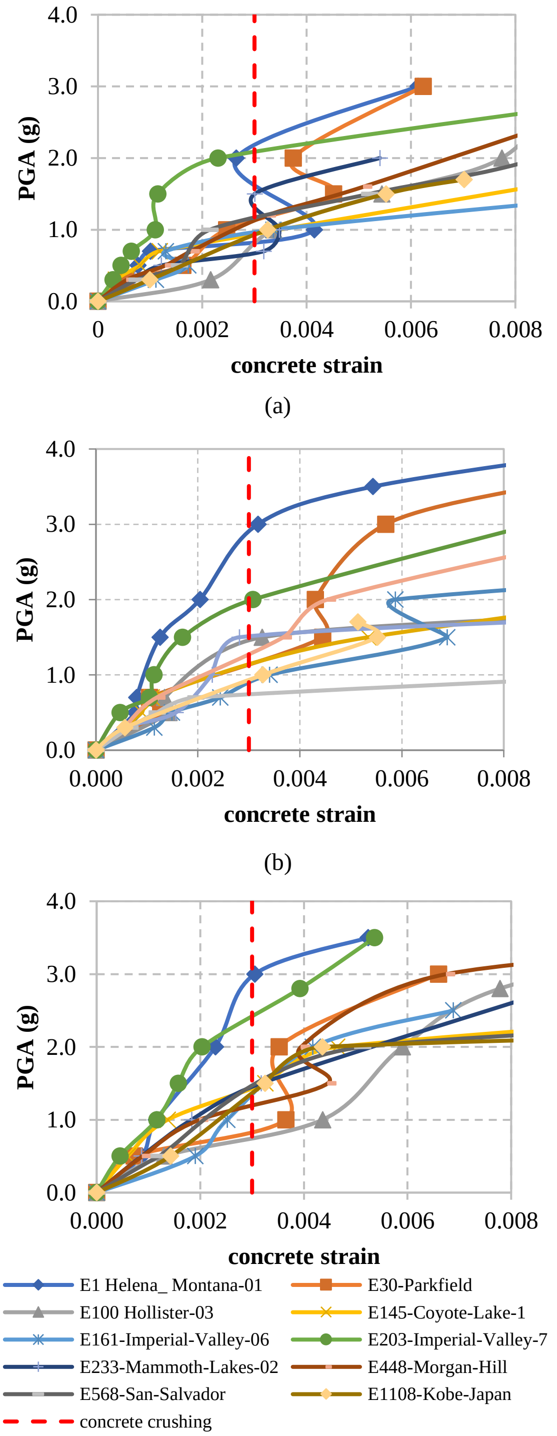


(a)

(b)

(b)

Figure 12. IDA curves for tensile strain of longitudinal reinforcements column with different FRC material: a) SFRC; b) PFRC; and c) HyFRC



(a)

(b)

(c)

Figure 13. IDA curves for concrete compressive strain of FRC columns material: a) PFRC; b) SFRC; and c) HyFRC

cumulative distribution, while c and β represent the median and standard deviation of the fragility curves, respectively. The estimation of the parameters c and β can be performed using the approach proposed by Baker (35). The damage states considered in this study were defined in section 2. The IDA curves based on the maximum drift ratio were utilized to generate the seismic fragility curves for the FRC columns at four limit states. Figure 14 illustrates that the HyFRC column exhibits a lower probability of exceeding the damage limit compared to the SFRC and PFRC columns. This indicates that the combination of fibers in RC columns effectively enhances the seismic performance. Moreover, the PFRC and SFRC columns show nearly identical seismic responses for extensive and complete damage states, while the SFRC and HyFRC columns demonstrate a better performance for moderate damage states.

4. RESULTS AND DISCUSSION

4. 1. Nonlinear Static Pushover Analysis

The analysis of the pushover curves demonstrates a significant improvement for the FRC column in term of shear capacity compared to RC columns; The HyFRC column exhibit the most significant improvement in seismic capacity, while SFRC and PFRC column ranked second and third. Table 6 illustrates the difference of displacement and base shear at four flexural damage states from the pushover curves. Table 6 provides a detailed breakdown of the disparities in displacement and base shear at four distinct flexural damage states as derived from the pushover curves. For steel yielding damage state, the PFRC is improved by 32 %, the SFRC by 63% and HyFRC by 75%. While for steel bar buckling, the PFRC is improved by 1 %, the SFRC by 17 % and HyFRC by 47 %. For steel bar fracture, the PFRC is improved by 10 %, the SFRC by 27 % and HyFRC by 58 %.

4. 2. Incremental Dynamic Analysis

By comparing IDA curves in terms of the maximum drift ratio, it can be concluded that the HyFRC columns experience lower drift ratios than the PFRC and HyFRC columns. Given the design PGA = 1.0 g at high seismic zones, the following maximum drift values are obtained from the fractile IDA curves:

- The 20% IDA curve records maximum drift values of 0.69%, 0.73%, and 0.79% for PFRC, SFRC, and HyFRC, respectively.
- The 50% IDA curve results in maximum drift values of 1.39%, 1.33%, and 1.27% for PFRC, SFRC, and HyFRC, respectively.
- The 80% IDA curve indicates maximum drift values of 1.60%, 1.44%, and 1.43% for PFRC, SFRC, and HyFRC, respectively.

In summary, it is evident that HyFRC is more effective in enhancing seismic capacity and reducing seismic demands compared to SFRC and PFRC.

4. 3. Seismic Fragility Analysis

In this study, we consider four predefined damage states: slight, moderate, extensive, and collapse, as previously outlined. Figure 14 presents the seismic fragility curves for the three FRC columns and the four damage states. Overall, there is a clear trend of decreasing damage probability as the damage state becomes more severe. The HyFRC columns are less fragile than the ones where the columns are made of SFRC and PFRC.

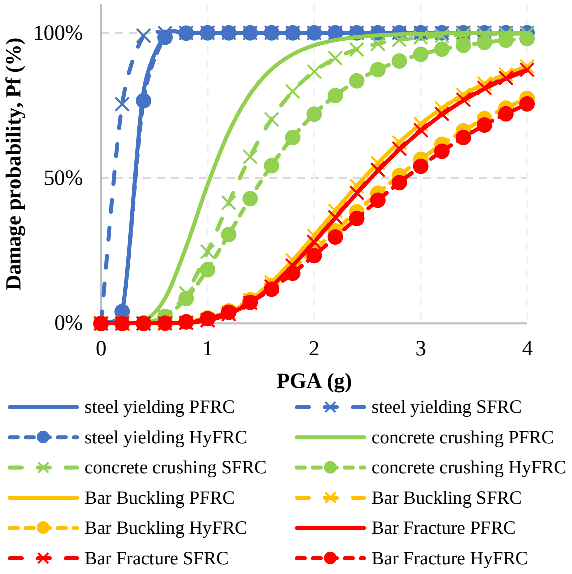


Figure 14. Seismic fragility curves of different FRC columns for four Performance levels

For the slight damage state, both HyFRC and PFRC demonstrate similar vulnerability. For instance, for example, the probability of damage exceedance at 0.2 g PGA is reduced from 75% for SFRC, to 4% for both PFRC and HyFRC.

At the moderate damage state, HyFRC exhibits the lowest damage vulnerability, with the probability of exceedance (for 1.4 g PGA) decreasing to 42.9%, from 57.3% and 78.9% for SFRC and PFRC respectively.

In the extensive damage state, PFRC and SFRC exhibit similar fragility while HyFRC is the least fragile. For a 2.6 g PGA, the probability of exceedance for PFRC and SFRC 44.9%, and 38.4% for HyFRC.

These findings highlight the effectiveness of a hybrid combination of steel and Polypropylene fibers with appropriate proportions in enhancing seismic performance compared to single-type of FRC columns. This is in agreement with previous studies (36, 37).

5. CONCLUSION

This study proposed a combination of CDP and XFEM to investigate the seismic performance of conventional reinforced concrete (RC) and fiber-reinforced concrete (FRC) columns. A 3D finite element model was developed for various FRC columns using appropriate constitutive laws and the XFEM technique to simulate cracking behavior. The model was calibrated using experimental data.

Nonlinear static pushover analysis (NSPA) was employed to evaluate the flexural damage states, and incremental dynamic analysis (IDA) was conducted on all FRC columns using a suite of 10 near-fault earthquake records. The maximum drift was selected as the engineering demand parameter (EDP) to compare the seismic behavior of different FRC columns. Based on the NSPA and IDA analysis results, the following conclusions can be drawn:

1. From the Pushover curves, it can be seen that there is a significant improvement for the FRC column in term of shear capacity compared to RC columns for all damage states.
2. By comparing IDA curves in terms of the maximum drift ratio, it can be concluded that the HyFRC columns experience lower drift ratios than the PFRC and HyFRC columns. Thus, it can be concluded that HyFRC are more effective to improve the seismic capacity and reduce the seismic demands of the column, compared to SFRC and PFRC.
3. The Seismic fragility analysis was conducted on the three FRC columns and the four damage states. The HyFRC columns are less fragile than the ones where the columns are made of SFRC and PFRC. Which highlights the effectiveness of a hybrid combination of steel and Polypropylene fibers in enhancing seismic performance compared to single-type of FRC columns.

6. REFERENCES

1. Mohammed AH, Mubarak HM, Hussein AK, Abulghafour TZ, Nassani DE. Punching Shear Characterization of Steel Fiber-Reinforced Concrete Flat Slabs. HighTech and Innovation Journal. 2022;3(4):483-90.
2. Bouzid L, Hamizi M, Hannachi N-E, Nekmouche A, Akkouche K. Plastic hinges mechano-reliability analysis in the beams of RC frames structures. World Journal of Engineering. 2020;17(5):719-32. 10.1108/WJE-02-2020-0069
3. Fanning PJ, Boothby TE. Three-dimensional modelling and full-scale testing of stone arch bridges. Computers & Structures. 2001;79(29-30):2645-62. 10.1016/S0045-7949(01)00109-2
4. Brencich A, Sabia D. Experimental identification of a multi-span masonry bridge: The Tanaro Bridge. Construction and Building Materials. 2008;22(10):2087-99. 10.1016/J.CONBUILDMAT.2007.07.031
5. Reccia E, Milani G, Cecchi A, Tralli A. Full 3D homogenization approach to investigate the behavior of masonry arch bridges: The Venice trans-lagoon railway bridge. Construction and Building Materials. 2014;66:567-86. 10.1016/J.CONBUILDMAT.2014.05.096
6. LI S, Yu T, Jia J. Empirical seismic vulnerability and damage of bottom frame seismic wall masonry structure: A case study in Dujiangyan (China) region. International Journal of Engineering, Transactions C: Aspects. 2019;32(9):1260-8. 10.5829/ije.2019.32.09c.05
7. Xiong Z, Chen J, Liu C, Li J, Li W. Bridge's Overall Structural Scheme Analysis in High Seismic Risk Permafrost Regions. Civil Engineering Journal. 2022;8(7):1316-27. 10.28991/CEJ-2022-08-07-01
8. 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, Transactions C: Aspects,. 2018;31(9):1505-13. 10.5829/ije.2018.31.09c.05
9. Conde B, Ramos LF, Oliveira DV, Riveiro B, Solla M. Structural assessment of masonry arch bridges by combination of non-destructive testing techniques and three-dimensional numerical modelling: Application to Vilanova bridge. Engineering Structures. 2017;148:621-38. 10.1016/J.ENGSTRUCT.2017.07.011
10. Cannizzaro F, Pantò B, Caddemi S, Caliò I. A Discrete Macro-Element Method (DMEM) for the nonlinear structural assessment of masonry arches. Engineering Structures. 2018;168:243-56. 10.1016/J.ENGSTRUCT.2018.04.006
11. Milani G, Lourenço PB. 3D non-linear behavior of masonry arch bridges. Computers & Structures. 2012;110:133-50. 10.1016/J.COMPSTRUC.2012.07.008
12. Oliveira DV, Lourenço PB, Lemos C. Geometric issues and ultimate load capacity of masonry arch bridges from the northwest Iberian Peninsula. Engineering Structures. 2010;32(12):3955-65. 10.1016/J.ENGSTRUCT.2010.09.006
13. Asadi P, Sourani H. Fragility curves production by seismic improvement of the high-dimensional model representation method. Engineering Computations. 2020;37(1):120-43. 10.1108/EC-12-2018-0586
14. Zhang Y, Macorini L, Izzuddin BA. Numerical investigation of arches in brick-masonry bridges. Structure and Infrastructure Engineering. 2018;14(1):14-32. 10.1080/15732479.2017.1324883
15. Baharmast H, Razmyan S, Yazdani A. Approximate incremental dynamic analysis using reduction of ground motion records. International Journal of Engineering, Transactions B: Applications. 2015;28(2):190-7. 10.5829/idosi.ije.2015.28.02b.04
16. Luco N, Cornell CA, editors. Effects of random connection fractures on the demands and reliability for a 3-story pre-Northridge SMRF structure. Proceedings of the 6th US national conference on earthquake engineering; 1998: Citeseer.
17. Vamvatsikos D, Cornell CA. Incremental dynamic analysis. Earthquake engineering & structural dynamics. 2002;31(3):491-514. 10.1002/eqe.141
18. Mohamed Nazri F, Kian Yern C, Moffed Kassem M, Noroozinejad Farsangi E. Assessment of structure-specific fragility curves for soft storey buildings implementing IDA and SPO approaches. International Journal of Engineering, Transactions C: Aspects,. 2018;31(12):2016-21. 10.5829/ije.2018.31.12c.04
19. Singh V, Sangle K. Analysis of vertically oriented coupled shear wall interconnected with coupling beams. HighTech and Innovation Journal. 2022;3(2):230-42. 10.28991/HIJ-2022-03-02-010
20. Zhang Y, Dias-da-Costa D. Seismic vulnerability of multi-span continuous girder bridges with steel fibre reinforced concrete columns. Engineering Structures. 2017;150:451-64. 10.1016/j.engstruct.2017.07.053
21. Huang L, Xu L, Chi Y, Xu H. Experimental investigation on the seismic performance of steel–polypropylene hybrid fiber reinforced concrete columns. Construction and Building Materials. 2015;87:16-27. 10.1016/j.conbuildmat.2015.03.073

22. Chi Y, Xu L, Zhang Y. Experimental study on hybrid fiber–reinforced concrete subjected to uniaxial compression. Journal of Materials in Civil Engineering. 2014;26(2):211-8. 10.1061/(ASCE)MT.1943-5533.0000764
23. Lubliner J, Oliver J, Oller S, Onate E. A plastic-damage model for concrete. International Journal of solids and structures. 1989;25(3):299-326. 10.1016/0020-7683(89)90050-4
24. Lee J, Fenves GL. Plastic-damage model for cyclic loading of concrete structures. Journal of engineering mechanics. 1998;124(8):892-900. 10.1061/(asce)0733-9399(1998)124:8(892)
25. Chi Y, Yu M, Huang L, Xu L. Finite element modeling of steel-polypropylene hybrid fiber reinforced concrete using modified concrete damaged plasticity. Engineering Structures. 2017;148:23-35. 10.1016/j.engstruct.2017.06.039
26. Abadel A, Abbas H, Almusallam T, Al-Salloum Y, Siddiqui N. Mechanical properties of hybrid fibre-reinforced concrete–analytical modelling and experimental behaviour. Magazine of Concrete Research. 2016;68(16):823-43. 10.1680/jmacr.15.00276
27. Hordijk D. Local Approach to Fatigue of Concrete Delft University of Technology. 1991.
28. Almusallam T, Ibrahim S, Al-Salloum Y, Abadel A, Abbas H. Analytical and experimental investigations on the fracture behavior of hybrid fiber reinforced concrete. Cement and Concrete Composites. 2016;74:201-17. 10.1016/j.cemconcomp.2016.10.002
29. Rousakis TC, Karabinis AI, Kiousis PD, Tepfers R. Analytical modelling of plastic behaviour of uniformly FRP confined concrete members. Composites Part B: Engineering. 2008;39(7-8):1104-13. 10.1016/J.COMPOSITESB.2008.05.001
30. Liang X, Xing P, Xu J. Experimental and numerical investigations of the seismic performance of columns with fiber-reinforced concrete in theplastic hinge region. Advances in Structural Engineering. 2016;19(9):1484-99. 10.1177/1369433216643896
31. Paulay T, Priestley MN. Seismic design of reinforced concrete and masonry buildings: Wiley New York; 1992.
32. Berry MP, Eberhard MO. Performance modeling strategies for modern reinforced concrete bridge. University of California, Berkeley. 2008.
33. Kowalsky MJ. Deformation limit states for circular reinforced concrete bridge columns. Journal of Structural Engineering. 2000;126(8):869-78. 10.1061/(ASCE)0733-9445(2000)126:8(869)
34. Liu M, Lu B, Liu B, editors. Study on performance index of reinforced concrete bridge column. Software Engineering and Knowledge Engineering: Theory and Practice: Volume 1; 2012: Springer. 10.4028/www.scientific.net/AMM.193-194.1079
35. Baker JW. Efficient analytical fragility function fitting using dynamic structural analysis. Earthquake Spectra. 2015;31(1):579-99. 10.1193/021113EQS025M
36. Pang Y, Cai L, Ouyang H, Zhou X. Seismic performance assessment of different fibers reinforced concrete columns using incremental dynamic analysis. Construction and Building Materials. 2019;203:241-57. 10.1016/j.cscm.2023.e02303
37. Zhang C, Cao M. Fiber synergy in multi-scale fiber-reinforced cementitious composites. Journal of Reinforced Plastics and Composites. 2014;33(9):862-74. 10.1177/0731684413514785

Persian Abstract

چکیده

اطمینان از مقاومت لرزه ای در مناطق زلزله خیز برای ایمنی سازه ضروری است. ستون های بتن تقویت شده با الیاف (FRC) نویدبخش بهبود عملکرد سازه در شرایط لرزه ای هستند. این مطالعه به دنبال ارزیابی جامع رفتار لرزه ای آنهاست. هدف اصلی این تحقیق ارزیابی و مقایسه عملکرد لرزه ای انواع ستون های FRC از جمله الیاف پلی پروپیلن (PFRC)، الیاف فولادی (SFRC) و ترکیبات ترکیبی (HyFRC) در مقایسه با ستون های بتن مسلح معمولی (RC) است. برای دستیابی به این هدف، این مطالعه از روش المان محدود توسعه‌یافته همراه با پلاستیسیته آسیب بتن (XFEM-CDP) در آباکوس برای بررسی دقیق پاسخ‌های استاتیکی و دینامیکی استفاده می‌کند. تجزیه و تحلیل فشار اور استاتیکی غیرخطی بهبود قابل توجهی در مقاومت لرزه ای در تمام انواع FRC در مقایسه با ستون های RC نشان داد. تحلیل‌های دینامیکی افزایشی (IDA) با استفاده از مجموعه انتخابی از ۱۰ حرکت زمین نزدیک به گسل برای ارزیابی پاسخ‌های لرزه‌ای غیرالاستیک ستون‌های پل FRC مختلف انجام می‌شود. شبیه‌سازی‌های XFEM-CDP در Abaqus جنبه‌های متعددی از ستون‌های FRC، مانند ترک‌خوردگی بتن، کاهش سختی و رفتار پلاستیک را به تصویر می‌کشند. تجزیه و تحلیل شکنندگی لرزه ای این ستون های FRC با در نظر گرفتن چهار حالت آسیب انجام می شود: الف) تسلیم طولی فولاد، ب) خرد شدن بتن هسته، ج) کمانش میله فولادی، و د) شکستگی میله فولادی طولی. نتایج نشان داد که ستون های HyFRC کمترین آسیب پذیری آسیب را در مقایسه با انواع PFRC و SFRC نشان می دهند.

IJE TRANSACTIONS B: Applications Vol. 37 No. 02, (February 2024) 283-297

International Journal of Engineering

Journal Homepage: www.ije.ir

Optical Efficiency of Linear Fresnel Reflectors in Fixed, Variable and Optimal Distance between Mirrors: Theoretical and Experimental Studies

A. Nakhaei Zadeh[a], M. Ameri*[a], A. Shojaei[a], I. Baniasad Askari[b]

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

PAPER INFO

Paper history:
Received 10 August 2023
Received in revised form 23 September 2023
Accepted 23 September 2023

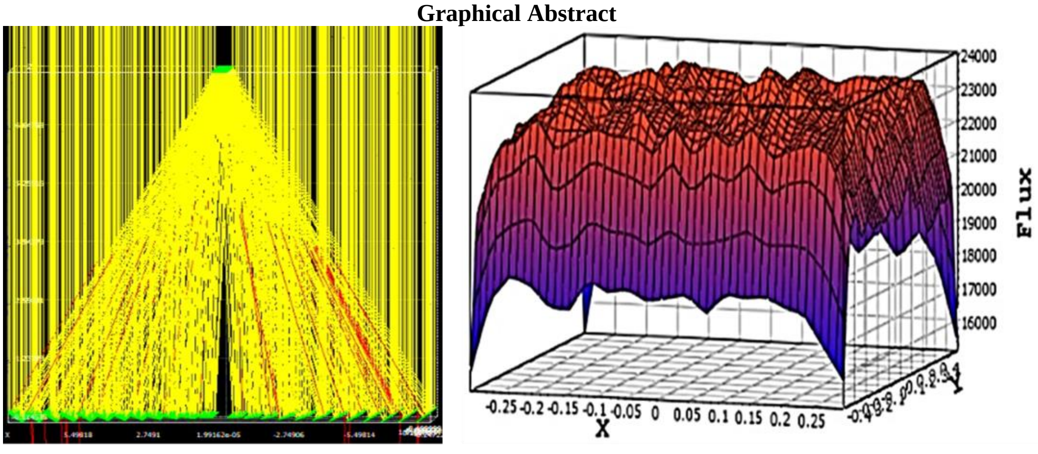
Keywords:
Linear Fresnel Reflectors
Solar Concentrators
Optical Efficiency
Thermal Efficiency
Optimization

ABSTRACT

In linear Fresnel reflectors, field arrangement has a significant effect on optical efficiency. Three constant, optimal, and variable distance layouts are proposed for Fresdemo solar power plant. The study was carried out by simulation and experiment. The small-scale Fresnel concentrator was designed and built with the capability to implement these three arrangements. The optical efficiency of the solar power plant with optimal variable and constant gap between mirrors were compared, considering fixed conditions for all layouts, including overall dimensions of the plant, width, and number of mirrors and dimensions of the receiver. It was observed that the arrangement with the optimal variable, and fixed distance had the highest to the lowest energy efficiency. Besides, the mirrors farther from the center entail more losses due to the sharper tilt angles, hence more spaces between these mirrors is required to reduce the losses. Meanwhile, the last mirrors in fixed distance arrangement have severe losses of shading and blocking while, they produce almost the same energy as central mirrors in the optimal and variable distance arrangement. The experimental results of the developed prototype showed that the thermal efficiency for the optimal distance was the highest, while it was followed closely by variable distance arrangement. The fixed distance arrangement had the lowest thermal efficiency. In addition, the variable and optimal distance arrangements exhibited an efficiency of 54% and 55%, respectively.

***doi**: 10.5829/ije.2024.37.02b.06*

Graphical Abstract



a) Ray-Tracing simulation b) Solar radiation distribution on the LFR receiver

---

*Corresponding Author Email: ameri_mm@uk.ac.ir (M. Ameri)

Please cite this article as: Nakhaei Zadeh A, Ameri M, Shojaei A, Baniasad Askari I. Optical Efficiency of Linear Fresnel Reflectors in Fixed, Variable and Optimal Distance between Mirrors: Theoretical and Experimental Studies, International Journal of Engineering, Transactions B: Applications. 2024;37(02):283-97.

NOMENCLATURE			
A_a	Absorber area (m^2)	q	Distance from the center of the field (m)
AFSC	Annular Fresnel solar concentrators	$\hat{S}$	Angle of the Sun's rays above the horizon on the transverse plane
AFSCF	Annular Fresnel solar concentrator coupled with a circular Fresnel lens	SPT	Solar Power Tower
B	Receiver's width (m)	s_n	Distance between two consecutive mirrors (m)
CSP	Concentrated Solar Power	T	Absolute temperature (K)
C_p	Specific heat at constant pressure (J/kg ℃)	T_i	Inlet temperature (K)
DNI	Direct Normal Irradiance, (W/m^2)	PTC	Parabolic Trough Collector
f	The height of absorber (m)	PDC	Parabolic Dishes Collector
f_b	The ratio of the blocked area of a mirror by the adjacent mirror to mirror surface	q	Distance from the center of the field (m)
f_s	The ratio of the shadowed area of a mirror by the adjacent mirror to mirror surface	T_o	Outlet temperature (K)
f_t	Total Loss Factor	T_a	Ambient temperature (K)
GA	Genetic algorithm	VDA	Variable distance arrangement
G_b	Beam radiation (W/m^2)	w_{nb}	The amount of unblocked area of a mirror (m^2)
LFC	Linear Fresnel Collector	w_{ns}	The amount of unshaded area of a mirror (m^2)
LFR	Linear Fresnel Reflector	**Greek Symbols**	
$\dot{m}$	Mass flow rate of fluid (kg/s)	$\widehat{\beta_\iota}$	Angle between the axis perpendicular to the mirror and horizon
$\widehat{M_l}$	Slope of the mirrors	$\widehat{\gamma_R}$	Angle of the line connecting the mirror center and receiver center with the horizon
ODA	Optimal distance arrangement	$\hat{\theta}$	Angle of the mirrors in the first generation of Fresnel collectors
PTC	Parabolic Trough Collector	η	Collector efficiency
PDC	Parabolic Dishes Collector	$\hat{\xi}$	Angle of inclination of the sun's rays

1. INTRODUCTION

The fast pace of technology and population growth have led to escalating CO$_2$ pollution, global warming, and fossil fuel depletion (1). Therefore, the use of renewable energy resources seems to be promising for electricity generation or water desalination (2, 3). Two different categories are usually defined for solar power plants, concentrating solar power plants (CSP) and photovoltaic systems (4). CSPs concentrate the solar energy on a fluid using different reflective instruments, leading to the generation of electrical energy based on the heated fluid in the turbine (5). The CPSs are divided into four categories concerning the geometrical characterization of reflectors, including parabolic dish collector (PDC), solar power tower (SPT) (6), parabolic trough collector (PTC), linear Fresnel collector (LFC) (7).

LFCs are a type of linear concentrating system with flat or slightly curved mirrors. The tilt of the mirrors changes to concentrate sunlight on the receiver in the center of the field. Two different generations have been developed for LFCs (8). In the first one, the fixed mirrors are arranged on a moving plate and the plate tracks the sun. In the second type, the mirrors reflect the sunlight on the receivers by changing their tilt. The first type became obsolete due to manufacturing difficulties and subsequent research was conducted on the second type. Losses of Fresnel collectors are divided into two categories: optical losses and thermal losses. The optical losses account for a larger share than thermal losses since these collectors usually are used at low and medium temperatures (9, 10). Fresnel systems are usually compared with linear parabolic collectors. A comparison of the advantages and disadvantages of LFCs and linear parabolic systems is summarized in Table 1.

The ease of construction and reduced costs of maintenance are the most critical advantages of linear Fresnel concentrators. Several parameters including the length and width of the mirrors, their distance from each other, receiver height, field position (north-south or east-west), field gradient, and number of mirrors are defined in the design process. The parameters not only provide designers with more freedom but also increase the complexity of design analysis and optimization. Therefore, these parameters need to be addressed in-depth to reduce losses in collectors and maximize collector efficiency based on Bellos et al. [9]. Every single parameter affects the rate of output energy. Assuming mirror distance from each other, mirror tilt, position of sun rays, and receiver height, some of the optical losses are detected in the system (see Table 2).

Several researchers have studied the linear Fresnel solar concentrators in both initial and secondary generations (11, 12). A comprehensive review in designing and application of the LFRs has been performed by Bellos et al. (13). In the study on the first generation of Fresnel solar concentrators conducted by Negi et al. (14), two relationships were suggested for the tilt and distance of mirrors in a Fresnel field. The tilt and distance of mirrors in a Fresnel field could be estimated by simultaneously resolving the mentioned relationships for conditions that the sun rays are normal to the collector. The zero shading and blocking losses, and maximizing the received power are the foundations of the mentioned relationships that were proposed for different receivers, including horizontal, vertical, and multi-tube receivers (15, 16). These relationships were applied only when the rays were normal to the moving plate; thus, the plate below the mirrors rotated such that it was always perpendicular to the sun rays. In other words, the models developed by this group were considered a type of solar island where the mirrors were fixed, and the plate below the mirrors was responsible for tracing the sun rays. They proposed some simple relationships to estimate the received energy and heat transfer (14). The second type of Fresnel concentrator is the new generation that has become the focus of new studies. The Solarmondo experimental power plant was built aimed at validating the ray-tracing of the sun's rays and investigating the corresponding functional and mechanical dimensions. Beltagy (17) studied the effect of glass on mono-tube and bi-tube receivers. In their study, a linear Fresnel prototype, developed by CNIM Company, was used in all tests. The optical field was made of 14 rows of mirrors with were constant distance between mirrors. The results showed that the removal of the receiver glass can increase annual optical efficiency by up to 5.6%. El Gharbi et al. (18)conducted a comparative study on the linear Fresnel solar concentrator with the east-west arrangement and linear parabolic concentrator, without considering the effects induced by shading or blocking the sun's rays. Abbas et al. (19) investigated the characteristics of concentrate in circular mirrors and determined the optimal width of mirrors based on the changed concentrate point in the mirror early and late in the day. Abbas and Martínez-Val (20) conducted an analytical study on the Fresdemo power plant in Spain. They investigated the effect of different parameters applying the ray-tracing methods.

In almost all research studies, the sun's rays are converted into transversal and longitudinal components. Zhu (21) proposed a vector analysis to determine the mentioned components. Ghodbane et al. (22) conducted a one-dimensional and transient numerical simulation on a LFR to evaluate the thermal efficiency, optical efficiency, and output temperature for the steam generator. That research was performed on 12 typical days during the year months for El-Oued, Algeria. Based on the results, the maximum thermal efficiency and optical efficiency were obtained as 37% and 53%, respectively. The curvature of mirrors was studied in some research. Lin et al. (23) constructed a prototype of concave mirrors with a bending radius of 3 m and an absorbent height of 1.5 m above the surface of the mirrors. There was a relatively good agreement between the results obtained by this group and the results reported by Morin et al. (24) where the bending radius was twice the distance between the mirror and the absorber. Heimsath et al. (25) investigated the effects induced by the degree of curvature, receiver height, and distance between mirrors using the ray-tracing methods. In their study, the north-south and east-west arrangements were also investigated. The north-south arrangement was finally considered a promising candidate. Santos et al. (26) performed the optical analysis of a LFR with a flat receiver considering the new analyses of shading, blocking, and cosine losses. The authors developed their analytical analysis in Phyton and compared the results with ray tracing simulations performed in SolTrace for both factorized and biaxial models of intercept factor. The results showed that the larger errors were obtained for high incidence angles ($\theta L \geq 80^\circ$). Sharma et al. (27) studied the compact linear Fresnel reflectors (LFRs) with two receivers on both sides of the optical field. They studied different optical field parameters such as the receiver's height, rows of reflectors, length and width of reflectors, and the effect of these parameters on electricity cost. Furthermore, Montes et al. (28) defined some new variables and criteria for analyzing a developed prototype. In their study, the threshold for radiation was 10 kw/h and this was the lowest value to be considered for heating fluid in the absorber. This study was performed for 3 optical fields with different widths and filling factors. Also, a comparison was done between Fresnel concentrators with single receiver and multi-receiver (compact linear Fresnel). It was claimed that although compact Fresnel concentrators have lower shading and blocking losses, these merits cannot outweigh the negative outcomes of the greater dispersion of sun rays (28). Roostaee and Ameri (29, 30) used an analytical method to investigate the dynamic relation between the reflector field and receiver optimum dimension for a constant width and shift arrangement in LFRs with a trapezoidal cavity receiver. They investigated the annual energy and exergy efficiency of Fresnel collectors. Kuchkarov et al. (31) developed a solar power plant module for individual consumers to obtain thermal, mechanical and electrical energy. An approach to determine the dimensions optimal diof flat mirrors was proposed based on the size of the receiver (27-32). Santos et al. (33) proposed a method to increase the amount of concentrated energy by increasing the number of primary concentrators (mirrors). In their research, the characteristics of the mirrors were the same in all the simulations, but with increasing the number of mirrors (total mirror area), the filling factor increased, which resulted in more energy concentration. Kincaida et al. (34) performed a sensitivity analysis on the insensitive parameters of the optical efficiency of a LFR to present insightful guidance for the manufacturing and implementation the optical losses. In that research, the

TABLE 2. Optical losses in the Fresnel concentrators

Loss type	Considerations
Shading	Adjacent mirrors cast shadows on each other[35].
Blocking	Reflection rays collide on the back of front mirrors in their path[30].
Irregular reflection of mirrors	Mirror surfaces do not have a perfectly regular reflection.
Error of solar tracking system	Mechanical system errors and limited accuracy of the device are the reason[24].
Receiver shadow	The receiver's shadow falls on each row of mirrors in a short time[35].

effect of twisting and Tracking errors of primary reflectors, the horizontal displacement of the receiver tube, and the vertical displacement of the secondary reflector on the LFR optical efficacy were investigated. A high-fidelity opto-mechanical error model was presented to obtain the realistic performance of the LFR. Said et al. (35) studied the numerical and experimental analysis of a small LFC considering the important parameters such as the longitudinal and transversal coefficients of the incidence angle modifier, the optical efficiency, and the concentration ratio. They concluded that the optical efficiency of the LFC significantly changes by the number of mirrors; 42.65% for 11 mirrors, 35.82%, and 26.98% for 9 and 7 mirrors, respectively. The previous studies on linear Fresnel concentrators have at least one of the following limitations:

- The existing Fresnel power plants have always been examined from a specific point of view and the effects of other parameters were eliminated (18).
- They tried to reduce the optical losses by defining some specific parameters.
- Optical field materials such as reflecting surface area and field factor were not constant (26, 30).

However, in the present study, the main contributions are as follows:

- Three arrangements with fixed, variable, and optimal spacing between mirrors were examined and compared for Kerman, Iran.
- The characteristics of mirrors, total mirror area, and filling factor were constantan for all arrangements, and it has been tried to achieve higher optical efficiency just by changing the place of mirrors. Also, the width of the field was the same for all three layouts, so that the center of the last mirror was a maximum of 8 meters from the center of the field, the absorber height was 8 m for all arrangements and the width of each mirror was 60 cm.
- It has been tried to achieve the highest optical efficiency by using the genetic algorithm in optimal distance arrangements. Moreover, by merging equations of the first generation of LFRs with the new generation, the variable distance arrangement obtained with the optical efficiency is close to the optimal arrangement.
- A small-scale prototype Fresnel concentrator was developed for practical testing of these three arrangements and comparing them. In the developed model, the location of the mirrors could be changed so that all the layouts could be implemented and checked. After several experiments on each arrangement, the efficiency versus loss parameter curve was finally plotted to compare the performance of these three arrangements in different environmental conditions.

2. NUMERICAL MODELS

In the initial studies on the Fresnel concentrators, all the mirrors were fixed on one plate whose function was to track the sun rays. Such a design requires a huge disc to hold all the mirrors along with a powerful guiding engine, also the disc movement led to receiver movement, highlighting the need to apply deformable connections on both sides of the receiver. The deformable connections led to the loss of the sealing costs and the chance of increasing the pressure up to high values. Changing the tilt of the mirrors in the next generation of Fresnel concentrators results in the concentration of sunlight on the receiver surface. In this design, the optical losses of Fresnel concentrators will be a concern. Section 2 deals with the governing equations related to these generations of Fresnel concentrators.

2. 1. The variable distance design

The mirror arrangement in the traditional design of the optical field was in a way that none of the mirrors had shading and blocking losses whenever the disc was normal to the sun rays. Although these concentrators are obsolete due to their disadvantages such as construction costs, huge sun tracker devices, and so on, their optical performance was extremely high, and this was a merit. Therefore, in this study, in the variable distance arrangement, the distance between the mirrors has been calculated from first-generation equations because in conditions where the sun rays are almost perpendicular to the surface (around solar

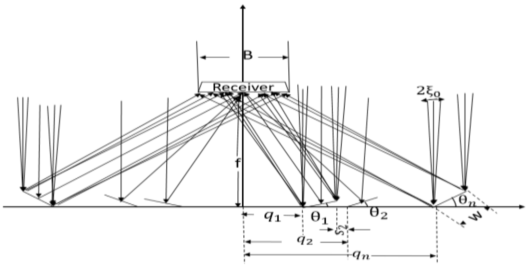


Figure 1. Transverse schematic of a Fresnel field (36)

noon), when the Direct Normal Irradiance (DNI) is also high, zero optical loss plays a vital role in optical performance. But the tilt of the mirrors is calculated according to the new generation relations in such a way that changing the tilted slope of each mirror causes the sun rays to focus on the receiver. In this way, construction difficulties in first-generation are eliminated. Furthermore, by tracking the sun rays during the day, concentrated energy is obtained. The relationships below are obtained for the first mirror using the geometry represented in Figure 1 (36).

$$q_1 = \frac{B}{2} + f \tan\left(\widehat{\xi_0}\right) \quad (1)$$

$$\widehat{\theta}_1 = \frac{1}{2}\left[tan^{-1}\left(\frac{q_1 + \frac{w}{2}\cos\left(\widehat{\theta_1}\right)}{f - \frac{w}{2}\left(sin\widehat{\theta_1}\right)}\right)\right] \quad (2)$$

where q_1 is the distance between the first mirror and center, θ_1 is the slope of the first mirror, ξ_0 is the angle of inclination of the sun rays, f is the receiver height, and B and w are the width of the receiver and mirrors, respectively. The first mirror distance from the field center is set as half the width of the receiver. The tilt is also set such that the rays radiated on the mirrors' center reach the receiver center. From the second mirror outwards, the relationships between location and rotation of mirrors are interrelated, and implicit equations should be solved simultaneously (36):

$$s_n = w\, sin\left(\widehat{\theta_{n-1}}\right) tan\left(2\widehat{\theta_n} + \widehat{\xi_0}\right) \quad (3)$$

$$q_n = q_{n-1} + w\, cos\left(\widehat{\theta_{n-1}}\right) + s_n \quad (4)$$

$$\widehat{\theta}_n = \frac{1}{2}\left[tan^{-1}\left(\frac{q_n + \frac{w}{2}cos\left(\widehat{\theta_n}\right)}{f - \frac{w}{2}sin\left(\widehat{\theta_n}\right)}\right)\right] \quad (5)$$

where, s_n is the distance between two consecutive mirrors. In this method, only the number of mirrors (n), width of every single mirror, and receiver height are required to design the field.

2. 2. The Constant Distance Design

The Fresnel concentrator mirrors are defined based on their location and the mirror's tilt angle at any given time. In this arrangement, the distance between mirrors is equal. According to Figure 2, the slope of mirrors changes to reflect rays radiated on the mirror center to the receiver center. The subsequent mirrors are placed in the position of $X_i + 1$ in priority so that the distance between each of them is defined as P_i. According to the mentioned law of reflection, the line perpendicular to each mirror at its point of rotation coincides with a bisector between radiated rays and the line connecting the mirror center with the receiver center on a transverse plane. Assuming that the vector perpendicular to the mirror creates a $\widehat{\beta}$ angle above the horizon, the equations below are obtained as follows (20):

$$\widehat{\beta_i} = \frac{\widehat{\gamma_R} + \hat{S}}{2} \quad (6)$$

$$\widehat{M_i} = \frac{\pi}{2} - \widehat{\beta_i} \quad (7)$$

where, $\widehat{M_i}$ is the slope of the mirror above the horizon at any given time. The mirrors are placed at an identical distance from each other concerning the filling factor of the field.

2. 3. Optimal field arrangement using genetic algorithm

A computer program was developed in MATLAB software for the optimization of the LF reflector using the Genetic Algorithm (GA) based on the pertinent input parameters. A genetic algorithm optimization problem is defined bellow (37):

- Minimizing or maximizing the objective function (or functions) included in the vector:

$$F(x) = f_1(x), f_2(x), \ldots, f_k(x)$$

- Satisfying the n restrictions of inequality and the m restrictions of equality:

$$\begin{cases} g_i(x) \geq 0 & i = 1, 2, \ldots, n \\ h_i(x) = 0 & i = 1, 2, \ldots, m \end{cases}$$

where, x is a vector which its elements are the decision variables of the problem.

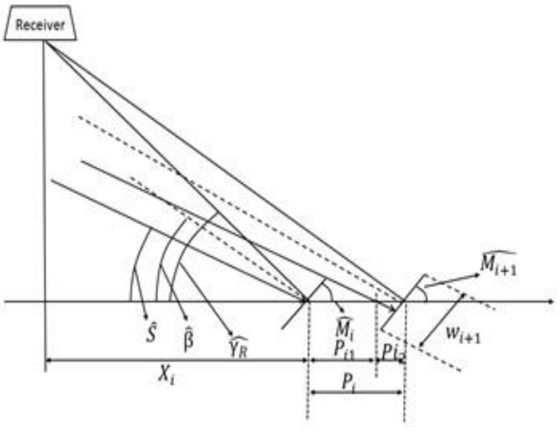


Figure 2. Radiation algorithm in a linear Fresnel concentrator (23)

- **Constrain bounds:**

$0.55\ cm \leq P_i \leq 8\ cm$

where, P_i is the distance between each two mirrors.

The optimization procedure of the GA is presented as follows:

1. Firstly, an initial population is randomly created.
2. Afterwards, a sequence of new populations is created by the algorithm by using the current generation individuals in the next population. The creation of the new population including the scoring of each population member by calculating its fitness value, scaling the raw fitness scores and converting them to the usable ranges, selecting the parents according to the fitness of each member, eliminating the less fitness members (elite), producing the children from the parents (using the mutation or cross over), and replacing the current population with the children to create the next generation.
3. Finally, the algorithm is terminated by applying the stopping criteria.

In this method, an arrangement is selected concerning the practical and spatial limitations, which results in the highest possible concentrated energy under the mentioned conditions. The objective function was annual concentrated energy $\left(F(x) = \sum_{j=1}^{j=8760} Q_{th}\right)$ and the goal was to maximize this objective function or minimize the reciprocal of this. The locations of the mirrors were considered as decision variables (f(x),), which included twelve mirrors on the right side of the field and twelve mirrors on the left side of the field symmetrically. The problem had geometric constraints including the dimensions of the field, so that the total width of the reflector is 255 cm and the distance between each two mirrors could change from 0.55 cm to 8 cm. Also, the location of the mirrors should not interfere with the process of tracking the sun. The genetic algorithm is depicted in Figure 3. In the first stage, some data such as solar data, power plant dimensions and receiver dimensions, and height are read. In the next step, the parameters of the genetic algorithm are determined. These parameters include mutation and crossover, number of iterations, and number and width of mirrors. The first population is then randomly constructed, and then optical losses such as shading of the mirrors and receiver, beam blocking, concentrated energy, and the objective function are calculated. Also, at this stage, the geometric conditions governing the problem are applied and then the answers that do not satisfy the geometric conditions of the problem are removed by applying the λ factor. This cycle is repeated and finally, the decision variables that provide the best objective function are introduced.

2. 4. Adjacent Mirrors Shading

The shade created on the mirror depends on the mirror's location, its distance from the adjacent mirror, and its slope. In all examined cases, the axis of rotation of mirrors coincides with their centers at a horizontal surface. As shown in Figure 4 (23), the mirrors with a width of w and the slope of M are positioned in a way that they can reflect the light radiated on the mirror center, on the receiver center at any given time. The value of shadow created at a given time is determined concerning the two adjacent mirrors and geometric considerations.

According to Figure 4, $\hat{S}$ is the angle of the sun's rays above the horizon on the transverse plane, $\widehat{\beta_\iota}$ is the angle between the axis perpendicular to the mirror and horizon direction, and γ_R is the angle of the line connecting the mirror center and receiver center with the horizon. Considering the geometric relationships between them, the following relations will be obtained (20).

$$\widehat{\gamma_R} = tan^{-1}\left(\frac{f}{X_i}\right) \tag{8}$$

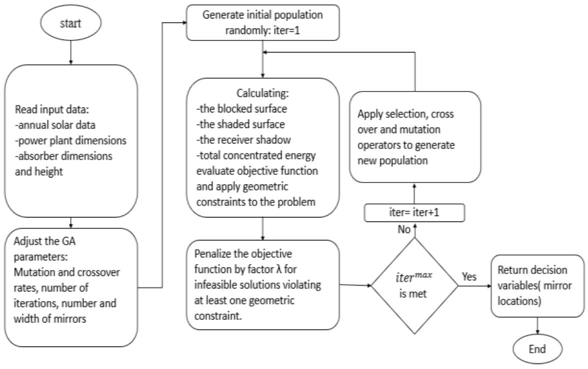


Figure 3. Genetic algorithm flow chart

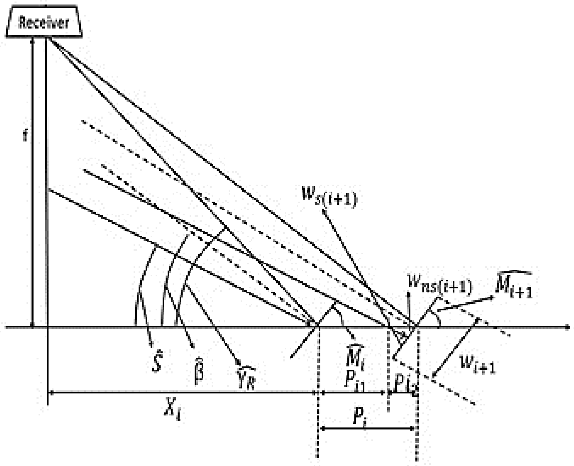


Figure 4. Shading in adjacent mirrors (23)

$$\widehat{\beta_\iota} = \frac{\widehat{\gamma_R}+\hat{S}}{2} \tag{9}$$

$$\widehat{M_\iota} = \frac{\pi}{2} - \widehat{\beta_\iota} \tag{10}$$

where, Equation 10 expresses the tilt of the mirror above the horizon at a given time.

By writing the sine relationships in the triangle specified in Figure 4, the following equation can be expressed (20, 36):

$$p_{i1} = \frac{w_i}{2} \cdot \frac{sin(\pi-\hat{S}-\widehat{M_\iota})}{sin(\hat{S})} \tag{11}$$

The law of sines is used to calculate the value of shadow created on the i+1th mirror. Accordingly, Equation 12 is obtained as follows (20, 36):

$$w_{ns(i+1)} = p_{i2} \cdot \frac{sin(\hat{S})}{sin(\pi-\hat{S}-\widehat{M_{\iota+1}})} \tag{12}$$

where $w_{ns(i+1)}$ is a portion of the $i+1^{th}$ mirror that has no shade.

By substituting $p_{i2} = p_i - p_{i1}$ and Equation 11 into above relation, Equation 13 is expressed as follows (20, 36) [23, 39]:

$$w_{ns(i+1)} = min\left[\left(p_i - \frac{w_i}{2} \cdot \frac{sin\ (\pi-\hat{S}-\widehat{M_\iota})}{sin\ (\hat{S})}\right) \cdot \frac{sin\ (\hat{S})}{sin\ (\pi-\hat{S}-\widehat{M_{\iota+1}})}, \frac{w_{i+1}}{2}\right] \tag{13}$$

The negative value of $w_{ns(i+1)}$ means that the created shadow extends beyond the mirror midpoint. The variable $f_{s(i+1)}$ expresses the ratio of shadowless portion to the surface of the mirror, which is defined based on Equation 14 and varies from zero to one under different conditions (20, 36):

$$f_{s(i+1)} = \frac{\frac{w_{i+1}}{2} - w_{ns(i+1)}}{w_{i+1}} \tag{14}$$

Given the equation defined for shadow, the maximum shadow occurs when $\pi - \hat{S} - \widehat{M_\iota} = \frac{\pi}{2}$, this refers to when the sun, centers of mirrors, and receivers are placed on a single line.

2. 5. Blocking in Adjacent Mirrors

Blocking is the incidence of reflected rays on the back of adjacent mirrors on the way to the receiver, which is another factor that is considered in the losses occurred in the Fresnel concentrators. Figure 5 represents the blocking process in the adjacent mirrors. The method taken to define the blocking relation is almost the same as the shading case (20, 36).

$$w_{nb(i+1)} = min\left[\left(p_i - \frac{w_i}{2} \cdot \frac{sin\ (\pi-\widehat{\gamma_{R_{\iota+1}}}-\widehat{M_\iota})}{sin\ (\widehat{\gamma_{R_{\iota+1}}})}\right) \cdot \frac{sin\ (\widehat{\gamma_{R_{\iota+1}}})}{sin\ (\pi-\widehat{\gamma_{R_{\iota+1}}}-\widehat{M_{\iota+1}})}, \frac{w_{i+1}}{2}\right] \tag{15}$$

where, w_{nb} indicates the portion of the mirror that delivers the reflected rays to the receiver. According to Equation 15, the maximum blocking occurs when $\pi - \widehat{\gamma_{R_{\iota+1}}} - \widehat{M_\iota} = \frac{\pi}{2}$, this means when the mirror is perpendicular to the line connecting the centers of the mirror and receiver. The variable $f_{b(i+1)}$ expresses the ratio of blocked surface to the total surface area of the mirror, which is defined based on Equation16 and varies from zero to one under different conditions (20, 36).

$$\mathrm{f_{b(i+1)}} = \frac{\frac{\mathrm{w_{i+1}}}{2} - \mathrm{w_{nb(i+1)}}}{\mathrm{w_{i+1}}} \tag{16}$$

2. 6. Overall Loss Coefficient of Adjacent Mirrors

The shading and blocking phenomena begin from the bottom of the mirror and then reach the mirror center or even beyond, consequently, when a mirror is simultaneously subjected to shading and blocking phenomena, the one with a more significant coefficient is predominant. These two factors should not be added.

$$f_t = max\ (f_s, f_b) \tag{17}$$

where f_t is the overall loss factor. Then, the optical efficiency of LFR could be obtained as follows (38):

$$\eta_{opt} = \alpha \times \rho_m \times \tau \times (1 - f_t) \tag{18}$$

where, α , ρ_m, and τ are the absorption factor of the receiver tube (0.94), mirror reflectance coefficient (0.92), and Cover glass transmittance (0.95), respectively.

3. EXPERIMENTAL AND SIMULATION SETUPS

3. 1. Measurement Instruments and Uncertainties

The measurement tools include the pyranometer, thermometer, flowmeter, and a potentiometer in the tracker system. A CM6B model pyranometer made of Kippzonen was used, which is capable of measuring the irradiation up to 2000 W/m^2 with a measurement uncertainty of 2%. Two thermometers, model PECULA, with stainless steel coils were used to measure the

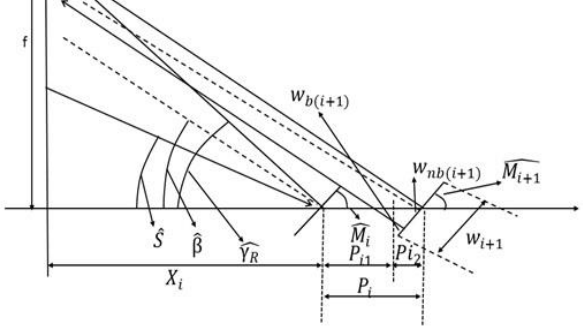


Figure 5. Blocking in adjacent mirrors (20)

temperatures at the inlet and outlet of the LFR absorber. These thermometers could measure the temperature within the range of 50-300\°C with a measurement uncertainty of 0.1\°C. In addition, a 10 kΩ potentiometer was applied in the tacker system which could report the rotational variations with a measurement uncertainty of 0.32 degree. Moreover, all the experiments were conducted under the open-loop conditions. So that the inlet water enters the LFR from the water reservoir with the mass flow rate of 60 L/h. The used flowmeter is model FTB-1424-HT, which measures the volumetric flow rate within the range of 20 L/h to 200 L/h with an uncertainty of 0.7% of the measured value.

3. 2. Design and Construction of a Prototype LFR

The solar collectors' thermal efficiency is determined either theoretically based on the optical and thermal properties of the constituent components or experimentally based on some given conditions. It should be noted that the theoretical thermal evaluation of collectors is accompanied by errors, which occurred at different boundary conditions. However, the theoretical thermal evaluation is always efficient in designing prototypes, and it is necessary to test the designed prototypes concerning different related standards. A small-scale Fresnel concentrator prototype was designed and constructed to compare constant, variable, and optimal distance arrangements. The linear Fresnel concentrator composed of 12 flat mirrors, six mirrors on each side, each with a thickness of 4 mm, width of 17 cm, and length of 180 cm with a total reflection surface area 3.67 m^2 was developed to evaluate the energy obtained in different field arrangements. Also, the total width of the LFR collector is 255 cm. The developed concentrator included three major parts: a concentrating field, a trapezoidal receiver, and a solar tracking system which is depicted in Figure 6. The trapezoidal receiver used in the concentrator had a large base of 28 cm, a small base of 12.8 cm, and a height of 12.5 cm. Three absorption vacuum tubes were used inside the receiver, each with a glass diameter of 58 mm and a steel tube diameter of 42 mm. The vacuum tubes were connected in series such that the working fluid (water) finally exited from the last tube after passing through three return paths. The input and output temperatures were measured using two thermometers with an accuracy of ±0.1 °C.

Mirrors were capable of moving along the frame, so by changing the location of the mirrors, it was possible to implement 3 different layouts. An 8 cm long shaft was located inside a bearing at the end of the frame of every single mirror to rotate the mirrors in order to track the sun’s rays at different times. Gears with a module of 2 and 18 teeth with a diameter of 4 cm were embedded at the end of the shaft. The teeth are in contact with a rack above them, which is driven by a linear actuator. The linear actuator is commanded using a control circuit programmed based on the solar data. Feedback seemed necessary due to changes in the speed of the actuator. To do this, a potentiometer of 10 kΩ was used to report the angular changes with an accuracy of 0.3268 degrees. Figure 7 shows the device's solar tracking system.

3. 2. Field Simulation with Ray-tracing Tools

Due to a large number of design parameters and system complexity in solar concentrator systems, it is impossible to physically examine or manufacture numerous practical and economic prototypes. Consequently, using computer codes not only leads to cost savings but also makes it possible to evaluate very complicated systems. Monte Carlo is a statistical method in which the random rays are generated based on the statistical methods. Such rays are generated at different radiation angles, and then the path related to every single ray is traced until it reaches the receiver or is lost in space. In the present study, the SolTrace software was used for simulating the optical field, which uses the Monte Carlo method for tracing the rays [42, 43]. Moreover, the optical properties of the materials used in the simulation of the LFR using SolTrace software are as shown in Table 3.

4. RESULTS

The data reported by Abbas and Martínez[23] on the losses calculated per mirror for Fresdemo power plant were used as reference data for validation of the method

Figure 6. The developed prototype

Figure 7. The developed solar tracking system

used in the present work. Firstly, the Fresdemo power plant in Almeria was considered in validation, in which 12 mirrors were symmetrically arranged on each side of the field with a filling factor of 71.4. According to Figure 8, there is a good agreement between the simulation results and those reported by Abbas and Martinez (20) indicating the accuracy of the applied method.

Then the simulation was done for local data in Kerman, Iran (30°17′N 57°05′E). In the constant distance arrangements (CDA), the distance between the center of the mirrors was considered identical and equal to 67 cm, while based on Equations 4 and 5 in the variable distance arrangement (VDA), the distance of the mirrors was calculated. In the optimal distance arrangements (ODA), the distances between mirrors were obtained using the genetic algorithm. The locations of 12 mirrors, which are the distances between the center of the mirrors and the center of the field on the right side of the field are presented in Table 4 for three arrangements. These distances are repeated exactly for the 12 mirrors on the left side.

The mirrors, which are farther from the center of the field, have more losses due to the sharper tilt angles during the day. Therefore, for the end mirrors, more distance is needed between these mirrors to minimize these losses. In variable and optimal distance arrangements, the distance between the mirrors gradually increases, which reduces optical losses while in constant distance arrangement, all distances between mirrors are equal. Figure 9 represents the shading and blocking in the first, sixth, and twelfth mirrors on the east side of the field in the constant, variable, and optimal distance arrangements concerning the receiver shading.

The Y-axis of Figures 9, 10, and 11 shows the total unblocked or unshaded mirror width which performs the reflection. The value of 0.6 m indicates that all mirrors with a total width of 0.6 m and without optical losses reflect the incident energy. The values lower than 0.6 m

TABLE 1. The optical properties of materials used in SolTrace

Properties	Value
Primary and secondary mirror reflectivity	0.95
Absorber tubes absorbance	0.94
External diameter of absorber tube	0.058 m
Internal diameter of absorber tube	0.042 m
Primary and secondary mirror transmissivity	0
Slope error	0.95 m rad
Specularity error	0.0001
Error type	Gaussian

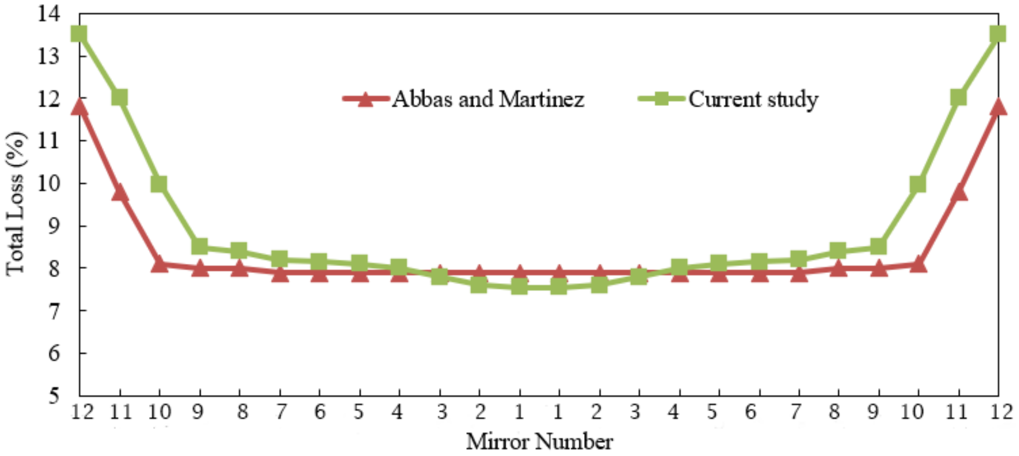


Figure 8. Comparison of percentage of energy losses in the field mirrors in the present study and the one conducted by Abbas and Martínez (20)

TABLE 4. Location of mirrors in the different field arrangements

Mirror number	1th	2th	3th	4th	5th	6th	7th	8th	9th	10th	11th	12th
						Mirrors distances						
CDA (m)	0.60	1.27	1.94	2.61	3.28	3.95	4.62	5.29	5.96	6.63	7.30	7.97
VDA (m)	0.63	1.24	1.84	2.46	3.09	3.73	4.38	5.06	5.76	6.48	7.23	7.98
ODA (m)	0.55	1.22	1.83	2.45	3.09	3.74	4.40	5.10	5.80	6.51	7.25	8.00

in Figures 9,10, and 11 show that only part of the width of the total mirrors participates in the reflection, thereby part of the incident energy is lost. According to the Figures 9,10, and 11 in the constant distance arrangement, just the central mirrors of the field concentrate more energy, while the last mirrors have severe losses of shading and blocking and have a small share in the total concentrated energy. However, mirrors with the optimal and variable distance arrangement generally operated better and produced more energy since the last mirrors produced almost the same energy as central mirrors of the field concentrate more energy, while the last mirrors have severe losses of shading and blocking and have a small share in the total concentrated energy. However, mirrors with the optimal and variable distance arrangement generally operated better and produced more energy since the last mirrors produced almost the same energy as central mirrors. The concentrated energy in the sixth mirror is almost the same for all arrangements because according to the curve gradients, shading has the greatest effect on the central mirrors, while blocking is more common on the end mirrors. Moreover, the performance of the last mirrors in the variable arrangement is very close to the optimal arrangement since the distance between mirrors gradually increases in these two arrangements.

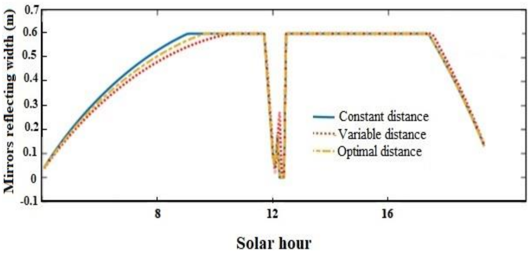


Figure 1. Shading or blocking on June 21 in the first mirror, numerical results

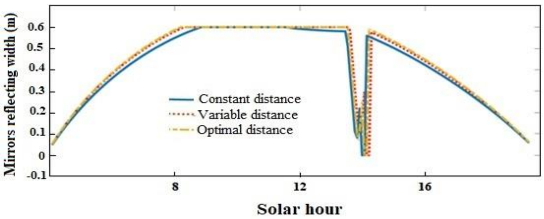


Figure 2. Shading or blocking on June 21 in the sixth mirror, numerical results

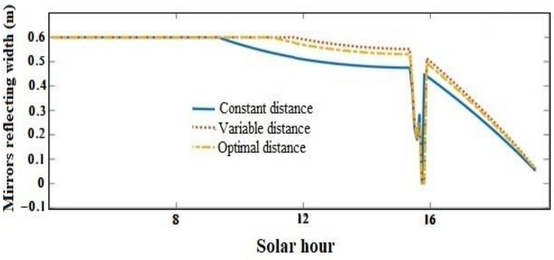


Figure 3. Shading or blocking on June 21 in the twelfth mirror, numerical results

The graph breaks show the effect of the receiver shadow. In some cases, the receiver's shadow overlaps with the losses caused by shading and blocking, and sometimes the receiver's shadow casts shadows on the parts that do not have the shading or blocking losses, in which case the maximum loss results.

By comparing the energy obtained in the various arrangements, it was concluded that the efficiency of concentrated energy in the optimal distance arrangement is about 1.3% and 0.8% higher than the constant distance and variable distance respectively. It should be mentioned that these values were calculated considering the effects induced by blocking and shading of mirrors and the receiver. Figure 10 represents the ratio of energy concentrated by each of the mirrors concerning different arrangements. According to Figure 12, while in the constant distance arrangement central mirrors have greater performance than the variable and optimal distance arrangements, the end mirrors are less involved in the concentrated energy. However, almost all mirrors are equally involved in the final output energy in the variable and optimal distance arrangement. This can be justified by the fact that in the variable and optimal distance arrangement, the effects induced by shading and blocking are reduced in the field edges as the distance between mirrors increases. In the constant distance arrangement, the end mirrors lose much of their reflective surface due to the effects induced by severe shading and blocking. The overall efficiency of the collector and its changes during the year are of special importance. Figure 13 shows the annual changes in the optical efficiency of the collector in three configurations.

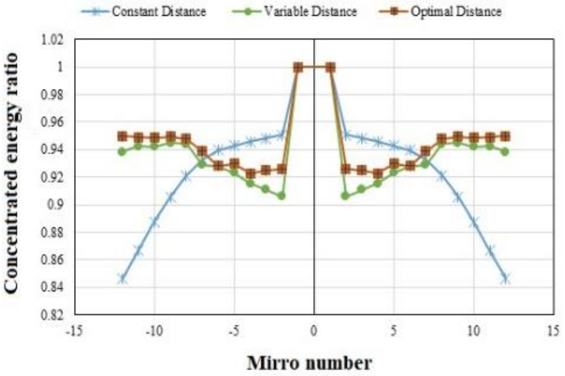


Figure 12. The mirrors energy shares in constant, variable and optimal distance arrangements, numerical results

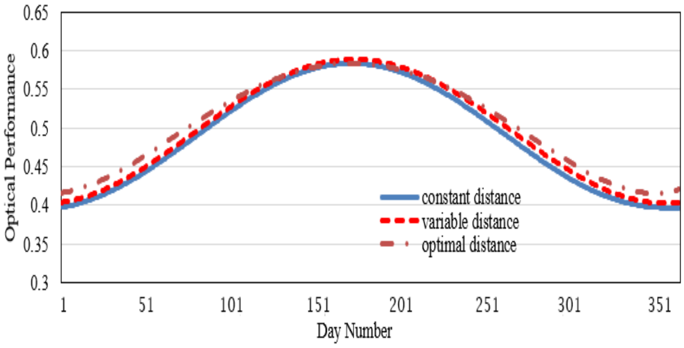


Figure 13. The yearly daily average optical efficiency in constant, variable and optimal distance arrangements, numerical results

According to Figure 13 all three arrangements have higher optical performance in the middle of the year than in the beginning and end of the year due to the angle of the sun and the losses caused by the transmission of the sun's rays to the mirror plates. Moreover, in the first and last days of the year, there is more difference in various layouts, while in the middle of the year, the optical efficiency of different layouts is almost the same. In general, the optical efficiency in the optimal distance arrangements is maximum and the fixed arrangement efficiency between the mirrors has the lowest efficiency, but in the middle of the year, the variable distance arrangement has a better performance than the optimal distance arrangement and different arrangement curves in some places are intertwined.

To assess the effect of beam spread, the constant and variable distance arrangements were separately modeled using the SolTrace software. It should be noted that in both layouts with variable and optimal spacing, the distance between mirrors gradually increases, and both layouts behave similarly in this regard, so only the results of the variable and constant distance arrangements are given. Figures 14 and 15 represent the concentration of energy in the constant and variable distance arrangements. The red rays indicate the reflections that do not reach the receiver. The optical profile of the back surface of mirrors is also defined as the front surface of mirrors to specify the blocked rays, which is why the red rays are reflected downwards. According to the figures, the lost rays occur mostly in the constant distance arrangement, particularly in the end mirrors. However, in the variable distance arrangement, deviations occur because of the mirrors' roughness and large distance of end mirrors.

Figures 16, 17 , 18 and 19 represent the flux distribution in the receiver for constant and variable distance arrangements.

As can be seen, because of using flat mirrors, the created flux is almost uniformly distributed on the receiver in both constant and variable distance arrangements, and there is no difference in the receiver center and its edges. In order to test the desired layouts, a small-scale prototype was designed in such a way that the location of the mirrors could be changed to test different arrangements. Each of the layouts was tested in May

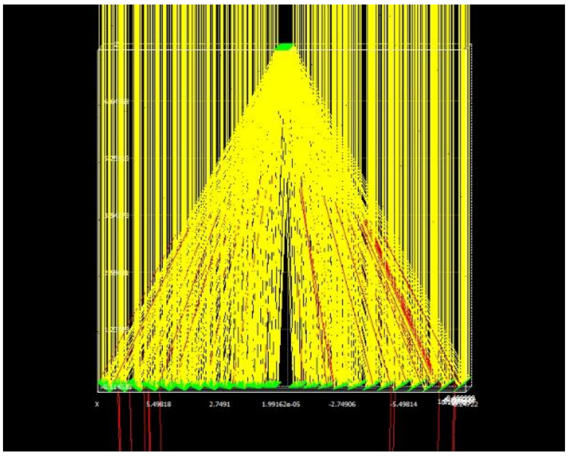


Figure 4. Energy concentration in the constant distance arrangement, simulation results

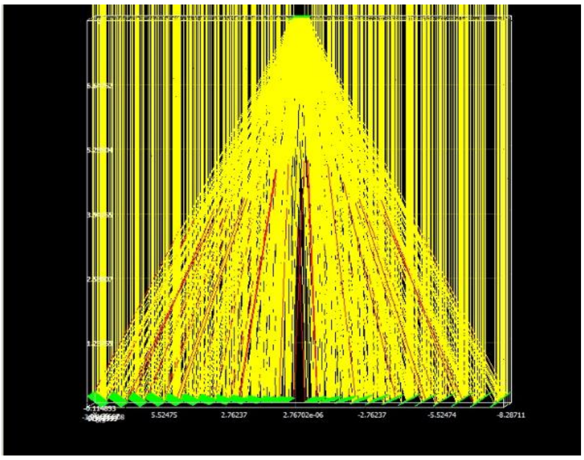
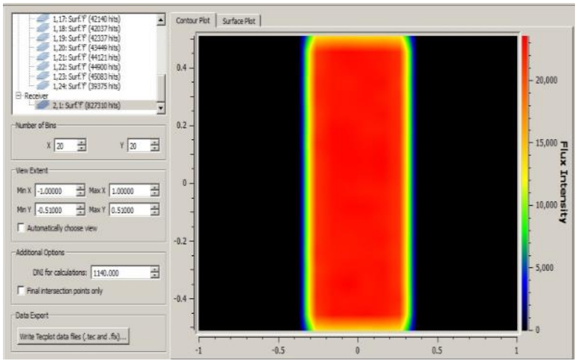
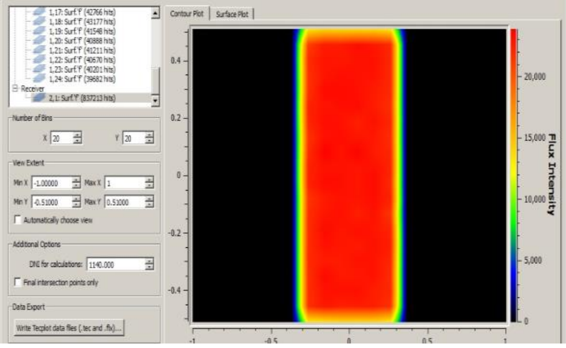


Figure 5. Energy concentration in the variable distance arrangement, simulation results


1,17: Surf.'f' (42140 hits)
1,18: Surf.'f' (42037 hits)
1,19: Surf.'f' (42337 hits)
1,20: Surf.'f' (43449 hits)
1,21: Surf.'f' (44121 hits)
1,22: Surf.'f' (44900 hits)
1,23: Surf.'f' (45083 hits)
1,24: Surf.'f' (39375 hits)
Receiver
2,1: Surf.'f' (827310 hits)

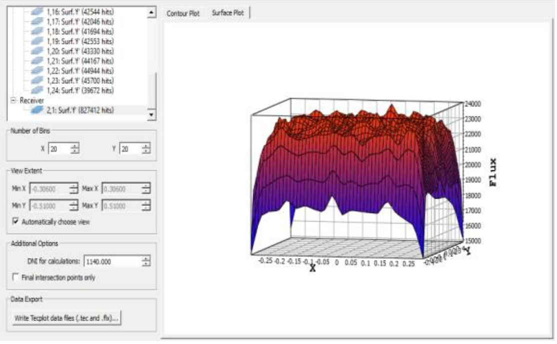
Number of Bins: X 20, Y 20
View Extent: Min X -1.00000, Max X 1.00000, Min Y -0.51000, Max Y 0.51000
☐ Automatically choose view
Additional Options: DNI for calculations: 1140.000
☐ Final intersection points only

Figure 6. The receiver flux contour in the constant distance arrangement, simulation results


1,17: Surf.'f' (42766 hits)
1,18: Surf.'f' (43177 hits)
1,19: Surf.'f' (41548 hits)
1,20: Surf.'f' (40888 hits)
1,21: Surf.'f' (41211 hits)
1,22: Surf.'f' (40670 hits)
1,23: Surf.'f' (40201 hits)
1,24: Surf.'f' (39682 hits)
Receiver
2,1: Surf.'f' (837213 hits)

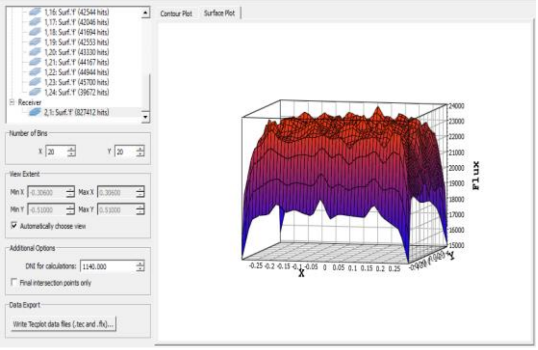
Number of Bins: X 20, Y 20
View Extent: Min X -1.00000, Max X 1, Min Y -0.51000, Max Y 0.51000
☐ Automatically choose view
Additional Options: DNI for calculations: 1140.000
☐ Final intersection points only

Figure 7. The receiver flux contour in the variable distance arrangement, simulation results


1,16: Surf.'f' (42544 hits)
1,17: Surf.'f' (42046 hits)
1,18: Surf.'f' (41694 hits)
1,19: Surf.'f' (42553 hits)
1,20: Surf.'f' (43330 hits)
1,21: Surf.'f' (44167 hits)
1,22: Surf.'f' (44944 hits)
1,23: Surf.'f' (45700 hits)
1,24: Surf.'f' (39672 hits)
Receiver
2,1: Surf.'f' (827412 hits)

Number of Bins: X 20, Y 20
View Extent: Min X -0.30600, Max X 0.30600, Min Y -0.51000, Max Y 0.51000
☑ Automatically choose view
Additional Options: DNI for calculations: 1140.000
☐ Final intersection points only

Figure 8. Surface distribution of radiation incident on the receiver in the constant distance arrangement, simulation results


1,16: Surf.'f' (42544 hits)
1,17: Surf.'f' (42046 hits)
1,18: Surf.'f' (41694 hits)
1,19: Surf.'f' (42553 hits)
1,20: Surf.'f' (43330 hits)
1,21: Surf.'f' (44167 hits)
1,22: Surf.'f' (44944 hits)
1,23: Surf.'f' (45700 hits)
1,24: Surf.'f' (39672 hits)
Receiver
2,1: Surf.'f' (827412 hits)

Number of Bins: X 20, Y 20
View Extent: Min X -0.30600, Max X 0.30600, Min Y -0.51000, Max Y 0.51000
☑ Automatically choose view
Additional Options: DNI for calculations: 1140.000
☐ Final intersection points only

Figure 9. Surface distribution of radiation incident on the receiver in the variable distance arrangement, simulation results

from 10 a.m. to 2 p.m. in Kerman (57.05 °E, 30.17 °N), while the time step was 15 minutes. Experiments were performed with north-south orientation in a closed loop, in which the following parameters were measured in each step: total solar radiation on the collector plate (G_t) diffuse solar radiation on the collector plane (G_d), ambient air temperature (T_a), Operating fluid inlet temperature (T_i), Operating fluid outlet temperature (T_o), and water flow $(\dot{m})$. By dividing the useful energy received by the input energy, the collector efficiency is obtained according to Equation 19.

$$\eta = \frac{\dot{m}c_p(T_o - T_i)}{A_a G_b} \tag{19}$$

In Equation 19 G_b is equal to the difference between G_t and G_d.

Graph η according to the loss parameter $(T_i - T_a)/G_b$ is drawn as a curve whose intersection with the vertical axis occurs when the inlet fluid temperature is equal to the ambient temperature and maximum efficiency occurs in these circumstances. The test results are shown in Figure 20. This figure represents the changes in the concentrator efficiency in terms of loss parameter. The specified points indicate the experimentally obtained points. The results obtained from the experimental tests are denoted by the small squares, triangles, and crosses in Figure 15 corresponding to the variable distance, optimal distance, and constant distance scenarios, respectively. By using a curve-fitting method (second-order polynomial fit), the curve that best fits the points are obtained for the LFR thermal efficiency (4), which are subsequently presented as Equations 20 to 22.

In the constant, variable and optimal distance arrangements, the equations of the second-degree curves are defined through Equations 20, 21 and 22, respectively.

$$\eta = 0.4503 - 0.0025\left(\frac{T_i - T_a}{G_b}\right) - (5.10^{-5})\frac{(T_i - T_a)^2}{Gb} \tag{20}$$

$$\eta = 0.5432 - 0.0028\left(\frac{T_i - T_a}{G_b}\right) - (6.10^{-5})\frac{(T_i - T_a)^2}{Gb} \tag{21}$$

$$\eta = 0.5577 - 0.0055\left(\frac{T_i - T_a}{G_b}\right) - (2.10^{-5})\frac{(T_i - T_a)^2}{Gb} \tag{22}$$

The constant distance arrangement has an efficiency of 45% under the best conditions, which is the least compared to the other two arrangements because, in constant arrangements, the last mirrors have great losses since there is not enough gap between them. Furthermore, central mirrors have unnecessarily large spaces and the sun rays should pass a larger distance to reach the receiver, this leads to larger spread of reflected rays; also, tracking system errors for first mirrors are

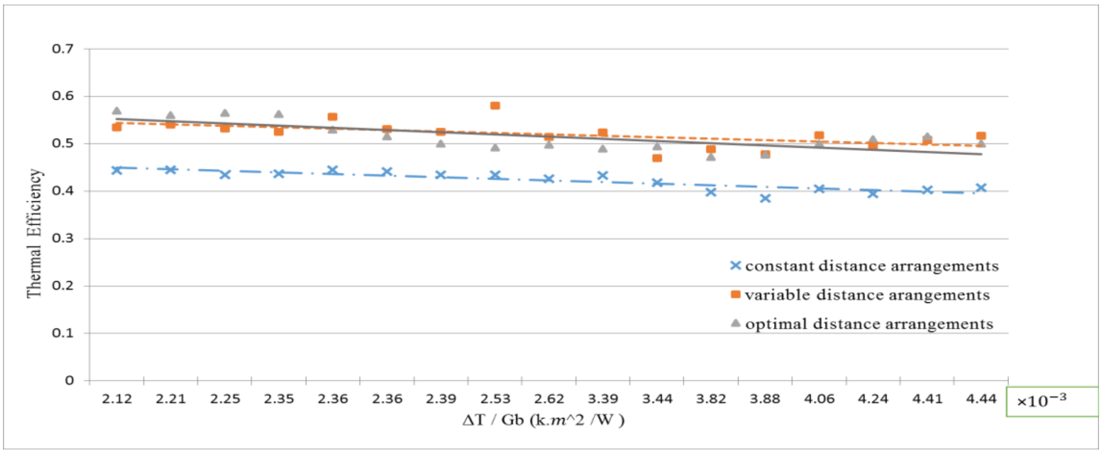


Figure 20. Thermal efficiency in terms of inlet water and ambient temperature difference obtained from the experimental tests

greater due to a larger distance between the mirrors and the receiver. In the results obtained from the experiment and simulation, the arrangement with the optimal distance between the mirrors has the highest efficiency, while the arrangement with the fixed distance between the mirrors has the lowest efficiency and a good agreement was observed between experiment and simulation. Since the width and the number of mirrors are small, the distance between mirrors in the optimal arrangement is mainly close to the variable arrangement so that they have almost the same performance, and their curves are intertwined in many points. Under the best conditions when the temperature of the working fluid is equal to the ambient, the variable and optimal distance arrangements have an efficiency of 54% and 55%, respectively. In these arrangements, the central mirrors are closer to each other, and these gaps are enough for the central mirrors to be out of the shading or blocking of the adjacent mirrors. The remaining space is used in the outer mirrors and they have greater distance in comparison with central mirrors. As a result, the field efficiency increases while the field width remains constant.

3. CONCLUSION

In the present study, three types of field arrangement with constant, variable, and optimal distance between mirrors was examined through a year while field dimensions were equal for the three arrangements. It is shown that:

To track the sun, the outer mirrors in the field always have a sharper angle than the central mirrors, and as a result, more optical losses occur in them.

- In the fixed-distance field layout, in which all mirrors have the same distance from each other, has large losses in the end mirrors, and these mirrors have a lower share of concentrated energy.
- In variable spacing, the distance between the mirrors gradually increases as they move toward the end mirrors, resulting in less loss in this arrangement than in fixed distance arrangement.
- In optimal field spacing, using the genetic algorithm and optimizing the amount of output energy, the location of each mirror was determined, and the simulation results indicated 1.2% and 0.8% efficiency improvement in comparison with constant distance and variable distance, respectively.
- The ray tracing simulation results showed that the end mirrors have a lot of optical loss compared to the central mirrors, and these losses had the highest value in the fixed distance arrangement and the lowest value in the optimal distance.
- The experimental results showed that the field arrangement at constant distance has a maximum efficiency of 45%, variable distance has a maximum efficiency of 54% and optimal layout has a maximum efficiency of 55%, so using genetic algorithms in the arrangement of the optical field of Fresnel concentrators can have a significant impact on the efficiency of these concentrators, although it does not increase costs. Moreover, the use of variable distance arrangements can be used as an efficient method in the design of Fresnel concentrators, which results in much better results than the fixed distance arrangement, and even on small scales has thermal efficiency close to optimal arrangement.

The future direction of the current study is to investigate the influence of using different nanofluids on the overall efficiency of the system.

4. REFERENCES

1. Sadeghi S, Baniasad Askari I. Performance and economic investigation of a combined phosphoric acid fuel cell/organic Rankine cycle/electrolyzer system for sulfuric acid production; Energy-based organic fluid selection. International Journal of Energy Research. 2020;44(4):2704-25. https://doi.org/10.1002/er.5073

2. Khaligh Fard S, Ahmadi H, Alizadeh Elizei MH. Electricity Supply Model of Conventional Residential Buildings in Tehran with Priority on Renewable Energy Using Adaptive Fuzzy-neural Inference System. International Journal of Engineering, Transactions A: Basics. 2023;36(10):1793-814. 10.5829/ije.2023.36.10a.07
3. Assari M, Mirzavand R, Basirat Tabrizi H, Jafar Gholi Beik A. Effect of steps height and glass cover angle on heat transfer performance for solar distillation: Numerical study. International Journal of Engineering, Transactions A: Basics. 2022;35(1):237-47. https://doi.org/10.5829/ije.2022.35.01A.23
4. Kalogirou SA. Solar energy engineering: processes and systems: Academic press; 2013.
5. Zhang H, Baeyens J, Degrève J, Cacères G. Concentrated solar power plants: Review and design methodology. Renewable and sustainable energy reviews. 2013;22:466-81. https://doi.org/10.1016/j.rser.2013.01.032
6. Sadeghi S, Askari IB. Parametric thermodynamic analysis and economic assessment of a novel solar heliostat-molten carbonate fuel cell system for electricity and fresh water production. Environmental Science and Pollution Research. 2022;29(4):5469-95. https://doi.org/10.1007/s11356-021-16035-2
7. Askari IB, Calise F, Vicidomini M. Design and comparative techno-economic analysis of two solar polygeneration systems applied for electricity, cooling and fresh water production. Energies. 2019;12(22):4401. https://doi.org/10.3390/en12224401
8. Negi B, Kandpal T, Mathur S. Designs and performance characteristics of a linear Fresnel reflector solar concentrator with a flat vertical absorber. Solar & wind technology. 1990;7(4):379-92. https://doi.org/10.1016/0741-983X(90)90023-U
9. Bellos E, Tzivanidis C, Moghimi M. Reducing the optical end losses of a linear Fresnel reflector using novel techniques. Solar Energy. 2019;186:247-56. https://doi.org/10.1016/j.solener.2019.05.020
10. Ma J, Chang Z, editors. Understanding the effects of end-loss on linear Fresnelcollectors. IOP Conference Series: Earth and Environmental Science; 2018: IOP Publishing. 10.1088/1755-1315/121/5/052052
11. Mokhtar G, Boussad B, Noureddine S. A linear Fresnel reflector as a solar system for heating water: theoretical and experimental study. Case Studies in Thermal Engineering. 2016;8:176-86. https://doi.org/10.1016/j.csite.2016.06.006
12. Said Z, Ghodbane M, Tiwari AK, Ali HM, Boumeddane B, Ali ZM. 4E (Energy, Exergy, Economic, and Environment) examination of a small LFR solar water heater: An experimental and numerical study. Case Studies in Thermal Engineering. 2021;27:101277. https://doi.org/10.1016/j.csite.2021.101277
13. Bellos E. Progress in the design and the applications of linear Fresnel reflectors–A critical review. Thermal Science and Engineering Progress. 2019;10:112-37. https://doi.org/10.1016/j.tsep.2019.01.014
14. Negi B, Mathur S, Kandpal T. Optical and thermal performance evaluation of a linear Fresnel reflector solar concentrator. Solar & wind technology. 1989;6(5):589-93. https://doi.org/10.1016/0741-983X(89)90095-7
15. Mathur S, Kandpal T, Negi B. Optical design and concentration characteristics of linear Fresnel reflector solar concentrators—I. Mirror elements of varying width. Energy Conversion and Management. 1991;31(3):205-19. https://doi.org/10.1016/0196-8904(91)90075-T
16. Mathur S, Kandpal T, Negi B. Optical design and concentration characteristics of linear Fresnel reflector solar concentrators—II. Mirror elements of equal width. Energy Conversion and Management. 1991;31(3):221-32. https://doi.org/10.1016/0196-8904(91)90076-U
17. Beltagy H. The effect of glass on the receiver and the use of two absorber tubes on optical performance of linear fresnel solar concentrators. Energy. 2021;224:120111. https://doi.org/10.1016/j.energy.2021.120111
18. El Gharbi N, Derbal H, Bouaichaoui S, Said N. A comparative study between parabolic trough collector and linear Fresnel reflector technologies. Energy Procedia. 2011;6:565-72. https://doi.org/10.1016/j.egypro.2011.05.065
19. Abbas R, Montes M, Piera M, Martínez-Val J. Solar radiation concentration features in Linear Fresnel Reflector arrays. Energy Conversion and Management. 2012;54(1):133-44. https://doi.org/10.1016/j.enconman.2011.10.010
20. Abbas R, Martínez-Val J. Analytic optical design of linear Fresnel collectors with variable widths and shifts of mirrors. Renewable Energy. 2015;75:81-92. https://doi.org/10.1016/j.renene.2014.09.029
21. Zhu G. Development of an analytical optical method for linear Fresnel collectors. Solar Energy. 2013;94:240-52. https://doi.org/10.1016/j.solener.2013.05.003
22. Ghodbane M, Boumeddane B, Said Z, Bellos E. A numerical simulation of a linear Fresnel solar reflector directed to produce steam for the power plant. Journal of cleaner production. 2019;231:494-508. https://doi.org/10.1016/j.jclepro.2019.05.201
23. Lin M, Sumathy K, Dai Y, Wang R, Chen Y. Experimental and theoretical analysis on a linear Fresnel reflector solar collector prototype with V-shaped cavity receiver. Applied Thermal Engineering. 2013;51(1-2):963-72. https://doi.org/10.1016/j.applthermaleng.2012.10.050.
24. Morin G, Dersch J, Platzer W, Eck M, Häberle A. Comparison of linear Fresnel and parabolic trough collector power plants. Solar energy. 2012;86(1):1-12. https://doi.org/10.1016/j.solener.2011.06.020
25. Heimsath A, Bern G, Van Rooyen D, Nitz P. Quantifying optical loss factors of small linear concentrating collectors for process heat application. Energy Procedia. 2014;48:77-86. https://doi.org/10.1016/j.egypro.2014.02.010
26. Santos AV, Canavarro D, Horta P, Collares-Pereira M. An analytical method for the optical analysis of Linear Fresnel Reflectors with a flat receiver. Solar Energy. 2021;227:203-16. https://doi.org/10.1016/j.solener.2021.08.085
27. Sharma V, Khanna S, Nayak JK, Kedare SB. Effects of shading and blocking in compact linear fresnel reflector field. Energy. 2016;94:633-53. https://doi.org/10.1016/j.energy.2015.10.098
28. Montes MJ, Rubbia C, Abbas R, Martínez-Val JM. A comparative analysis of configurations of linear Fresnel collectors for concentrating solar power. Energy. 2014;73:192-203. https://doi.org/10.1016/j.energy.2014.06.010
29. Roostaee A, Ameri M. Effect of Linear Fresnel Concentrators field key parameters on reflectors configuration, Trapezoidal Cavity Receiver dimension, and heat loss. Renewable Energy. 2019;134:1447-64. https://doi.org/10.1016/j.renene.2018.09.053
30. Roostaee A, Ameri M. A comparative study of different optimised mirrors layouts of Linear Fresnel concentrators on annual energy and exergy efficiencies. International Journal of Ambient Energy. 2022;43(1):2627-44.
31. Kuchkarov A, Abdumuminov A, Abdurakhmanov A. Developing a Design of a Composite Linear Fresnel Mirror Concentrating System. Applied Solar Energy. 2020;56:192-7. 10.3103/S0003701X20030056
32. Sharma V, Nayak JK, Kedare SB. Effects of shading and blocking in linear Fresnel reflector field. Solar Energy. 2015;113:114-38.
33. Santos AV, Canavarro D, Collares-Pereira M. The gap angle as a design criterion to determine the position of linear Fresnel

primary mirrors. Renewable Energy. 2021;163:1397-407. https://doi.org/10.1016/j.renene.2020.09.017

34. Kincaid N, Mungas G, Kramer N, Zhu G. Sensitivity analysis on optical performance of a novel linear Fresnel concentrating solar power collector. Solar Energy. 2019;180:383-90. https://doi.org/10.1016/j.solener.2019.01.054
35. Said Z, Ghodbane M, Hachicha AA, Boumeddane B. Optical performance assessment of a small experimental prototype of linear Fresnel reflector. Case Studies in Thermal Engineering. 2019;16:100541. https://doi.org/10.1016/j.csite.2019.100541
36. Mathur S, Negi B, Kandpal T. Geometrical designs and performance analysis of a linear Fresnel reflector solar concentrator with a flat horizontal absorber. International journal of energy research. 1990;14(1):107-24. https://doi.org/10.1002/er.4440140111
37. Goel A, Manik G. Step towards sustainability: Techno-economic optimization of a parabolic trough solar collector using multi-objective genetic algorithm. Thermal Science and Engineering Progress. 2023;37:101539. https://doi.org/10.1016/j.tsep.2022.101539
38. Rungasamy A, Craig K, Meyer JP. A review of linear Fresnel primary optical design methodologies. Solar Energy. 2021;224:833-54. https://doi.org/10.1016/j.solener.2021.06.021

Persian Abstract

چکیده

در بازتابنده های فرنل خطی، آرایش میدان تاثیر بسزایی بر بازده نوری دارد. سه طرح فاصله ثابت، بهینه و متغیر برای نیروگاه خورشیدی Fresdemo پیشنهاد شده است. مطالعه به صورت شبیه سازی و آزمایش انجام شده است. متمرکز کننده فرنل در مقیاس کوچک با قابلیت اجرای این سه ترتیب طراحی و ساخته شد. راندمان نوری نیروگاه خورشیدی با فاصله بهینه، متغیر و ثابت بین آینه ها با در نظر گرفتن شرایط ثابت برای همه چیدمان ها از جمله ابعاد کلی نیروگاه، عرض و تعداد آینه ها و ابعاد گیرنده مقایسه شده است. مشاهده شد که چیدمان با فاصله بهینه، متغیر و ثابت دارای بیشترین تا کمترین بازده انرژی است. علاوه بر این، آینه های دورتر از مرکز به دلیل زوایای شیب تیزتر تلفات بیشتری را به همراه دارند، بنابراین برای کاهش تلفات به فضاهای بیشتری بین این آینه ها نیاز است. این در حالی است که آخرین آینه ها در آرایش فاصله ثابت دارای تلفات شدید سایه و بلوکه شدن هستند در حالی که در آرایش فاصله بهینه و متغیر تقریباً انرژی معادل با آینه های مرکزی تولید می کنند. نتایج تجربی نمونه اولیه توسعه‌یافته نشان داد که راندمان حرارتی برای فاصله بهینه بالاترین بود، در حالی که چیدمان با فاصله متغیر بازدهی تقریبا مشابه داشت و چیدمان فاصله ثابت کمترین بازده حرارتی را داشت. علاوه بر این، چیدمان فاصله متغیر و بهینه به ترتیب بازدهی ٥٤٪ و ٥٥٪ را نشان دادند.

IJE TRANSACTIONS B: Applications Vol. 37 No. 02, (February 2024) 298-311

International Journal of Engineering

Journal Homepage: www.ije.ir

A Complex Network-based Approach for Designing of Wide Area Measurement Systems in Smart Grids using Adam-Eve Like Genetic Algorithm

M. Shahraeini*, R. Soltanifar

Department of Electrical Engineering, Faculty of Engineering, Golestan University, Gorgan, Iran

PAPER INFO

Paper history:
Received 07 August 2023
Received in revised form 21 September 2023
Accepted 23 September 2023

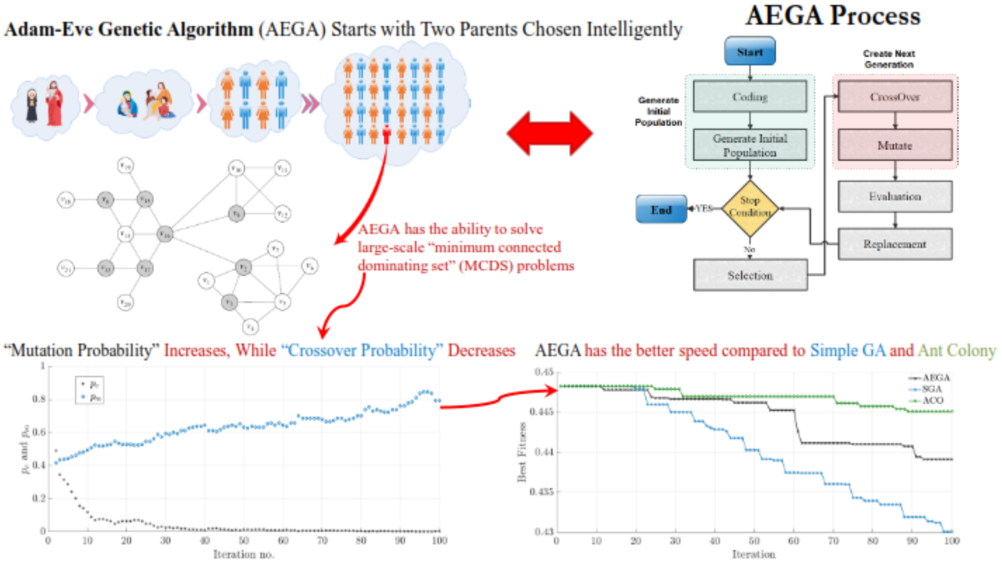
Keywords:
Wide Area Measurement System
Minimum Connected Dominating Set
Genetic Algorithm
Adaptive Genetic Algorithm
Ant Colony Optimization

ABSTRACT

The wide area measurement system (WAMS) consists of two different measuring and communication infrastructures, which is respectively responsible for measuring power girds' data in the wide area and sending and processing them in the control centers. The design of WAMS can include the design of each of its infrastructures or target both infrastructures at the same time, the latter has been known as the WAMS comprehensive design. The WAMS comprehensive design means the simultaneous placement of measurement components and its required communication, which is known as minimum connected dominating set (MCDS) problem in graph theory and is formulated in the form of an optimization problem. Solving such a complex optimization problem is often done with evolutionary algorithms (e.g. genetic algorithm and ant colony), and the speed and efficiency of finding the solution has always been a challenge. This research proposes an adaptive genetic algorithm known as the Adam and Eve algorithm, which has the ability to solve the MCDS problem that arises from the WAMS comprehensive design. Through simulation results for IEEE 1354 bus network, we demonstrate that proposed algorithm is well-tuned to solved MCDS related to the power graphs. It is 30% faster than simple genetic algorithm, handles large-scale problems effectively, and outperforms both simple genetic algorithm and ant colony algorithm within a given timeframe.

doi: *10.5829/ije.2024.37.02b.07*

Graphical Abstract



---

*Corresponding Author Email: *m.shahr@gu.ac.ir* (M. Shahraeini)

Please cite this article as: Shahraeini M, Soltanifar R. A Complex Network-based Approach for Designing of Wide Area Measurement Systems in Smart Grids using Adam-Eve Like Genetic Algorithm, International Journal of Engineering, Transactions B: Applications. 2024;37(02):298-311.

1. INTRODUCTION

The term “Wide Area Measurement System” (WAMS) was first introduced to the power system literature in the late 20th century to describe a novel measurement system that has become a crucial component of smart grids. Such systems include new advanced digital measurement devices (e.g. PMUs), and modern and high-speed communication infrastructure that enable the efficient management of the complex behavior demonstrated by extensive power grids. In general, the WAMS consists of three interconnected subsystems: data acquisition, data transmission, and data processing (1, 2).

Normally, “WAMS design” may concentrate on any of the mentioned subsystems. For example, placing of the PMUs is considered as the design of the measurement subsystem, while the design of the communication infrastructure for such units implies the design of the communication subsystem of the WAMS. The former is known as the optimal PMU placement (OPP) problem (3), while the latter can be considered as the optimal OPGW placement (OOP) problem (4). It should be noted that the PMU has the ability to measure the voltage and current phasor with high resolution in the entire of the power grid, and OPGW optical fiber is preferred as the transmission medium for these units due to its low latency and high capacity. It can be installed as a part of transmission lines (2).

Both above mentioned problems (i.e. OPP and OOP) can be formulated as the optimization problems (5). In WAMS design, there is an opportunity to place different WAMS components (i.e. PMUs and OPGWs) of two subsystems at the same time, which we call it “*WAMS comprehensive design*” (6). This can be also considered as optimal PMU and OPGW placement problem (OPOP) and it is also can be formulated as the optimization problem (5).

As explained above, different WAMS design problems may be formulated as the optimization problems and should be solved by optimization problem-solving methods. In general, there are three main categories of methods for solving optimization problems (7): deterministic, stochastic, and hybrid. Deterministic methods, also known as Gradient-based methods, are generally faster computationally and can find local optimum. Examples of deterministic methods include Gradient Conjugate, Newton, and Dayton-Fletcher-Powell methods (8). On the other hand, stochastic methods (also known as evolutionary algorithms) are slower but have the potential to find global optimum. Examples of stochastic methods include genetic algorithm (GA), simulated annealing (SA), ant colony (ACO), and particle swarming algorithms (PSO). To address the limitations of both deterministic and stochastic methods, hybrid methods have been developed that combine the robustness of stochastic methods with the speed of deterministic ones (7).

Among the evolutionary algorithms introduced above, the genetic algorithm is a widely used and highly effective evolutionary algorithm that can tackle various problems across different domains. It is also a good choice for multi-objective problems (9-11). GAs may be classified into several categories, including simple genetic algorithm, parallel and distributed genetic algorithm, fast messy genetic algorithm, adaptive genetic algorithm, hybrid genetic algorithm, and independent sampling genetic algorithm (12, 13).

The design problems discussed earlier (i.e. WAMS design problems) have equivalents in the complex networks, with the OPP referred to as the “minimum dominating set” (MDS) problem and the OPOP known as the “minimum connected dominating set” (MCDS) problem (5). These are both considered to be NP-hard problems (14), and their effectiveness is a major concern.

In particular, the MCDS problem is a more complex problem because it performs the simultaneous placement of two different categories of WAMS components (i.e., PMUs and OPGWs) and is often a multi-objective optimization problem with various constraints in which effectiveness becomes more important in such a problem (15-18). Besides the efficiency in solving of MCDS (hereafter we call it “*performance*”), the speed of obtaining the solution (hereafter we call it “*speed*”) is also a significant consideration, particularly when dealing with large problem sizes.

The complexity of the MCDS problem and the emphasis on performance and speed has led to the introduction of sequential techniques in addition to the comprehensive approach for solving the MCDS problem. Unlike the comprehensive approach that solves the entire MCDS problem in the form of an optimization problem, sequential methods first place PMUs and then, using complex network-based algorithms, OPGWs are placed to produce required WAMS communication infrastructure. In these methods i.e., sequential ones, the placement of PMUs is often in the form of optimization, while graph algorithms are generally based on the shortest path.

The problem of addressing the MCDS problem in smart grids and comprehensive designing of wide area measurement systems has been a concern for approximately a decade. For the first attempt made by Fan and Watson (19) and by using integer programming, were solved MCDS problem for simultaneous placement of PMUs and their required communication infrastructure but there; they didn’t point to MCDS as the WAMS comprehensive design. At the same time, the WAMS comprehensive design was firstly introduced by Shahraeini et al. (6), while the authors didn’t directly point to MCDS problem. There, simple genetic algorithm

(SGA) was used to solve proposed problem and the solution was of great importance, while performance and speed of SGA were not of main concern. After that, many researches concentrate on improvement of solutions for WAMS design, while many of them introduce sequential methods for WAMS design (15-18), that is, they first perform OPP and then place communication links by graph shortest path algorithms. Meanwhile, few studies (e.g. (20)) have focused on solving the MCDS problem for WAMS design, that is, they place PMUs and OPGWs in one run and by solving MCDS problem. In our recent works, we have introduced ACO for solving MCDS problem with some heuristics such as pheromone modification. There, we have dramatically improved performance compared to SGA results presented by Shahraeini et al. [6], but still the speed of reaching to the best solution has been not improved there. As a new approach, machine learning and deep learning algorithms (e.g. methods presented in literature (21-27)) are another approach to study MCDS; for instance, convolutional neural networks (CNNs) can be used to classify nodes in a graph as either part of the MCDS or not.

A review of previous works presented above indicates that due to NP-hardness of MCDS problem, performance and speed of methods used for solving of such a problem (either comprehensive methods or sequential ones) are still of main concern. On the other hands, stochastic method like GA and ACO are preferred methods in solving MCDS as the comprehensive approach since they are simple to implement and at the same time guarantee to find at least one local optimum; but they have been not examined yet in large-scale problems for the large-scale power grids. As a result of these facts, introducing new stochastic methods for improvement of performance and speed in solving MCDS problems that has the ability to solve such large-scale problems are still an opening issue for researchers and this is the main motivation of this study.

The main objective of this study is to propose an adaptive genetic algorithm to solve MCDS problem that is derived from WAMS comprehensive design. In our very last publication (28), we have introduced Adam-Eve Genetic Algorithm (AEGA) as an adaptive GA to solve OPP problem. There, we have shown the advantages of AEGA over other GA algorithms like SGA. In this study, we have designed previously proposed AEGA (28) to solve MCDS (i.e. OPOP) problem in large-scale power grids. To the best of our knowledge, this is the first attempt to implement an adaptive GA to solve MCDS problem in power graphs, and also the first study that addresses the solving of large-scale problems.

The main contributions of this paper are as follows:

- The design and setting of the parameters of an adaptive GA (e.g. AEGA) strongly depends on the nature and type of the problem and is inherently hard, which requires appropriate heuristics such as choosing initial population and well-tuning of crossover and mutation rates. In this study and for the first time, this is well done for the MCDS problem raised from large-scale power grids.
- Although MCDS problem is an offline design problem, algorithms such as ACO and SGA, whose high performance has been proven before [5-6], are unable to solve such large-scale problems in a reasonable amount of time. For this reason, the speed of reaching to the solution is very important in large-scale power grids, and the proposed AEGA has improved the speed by 30% compared to its basic version, i.e., SGA.
- The speed improvement in OPOP that has been done in the current research, along with the performance improvement in OPP that was done in our previous research (28), is a proof of the fact that adaptive GAs have a good ability to improve solving of the problems related to complex networks such and MDS and MCDS problems. Additionally, better performance during the same execution time than SGA and ACO in large-scale power graphs is another advantage of the proposed AEGA that has been shown in the current research.
- Choosing Adam and Eve at the beginning of AEGA that is proposed in this study makes the search space closer to the optimal solutions, and this is a considerable improvement in performance and speed. We have shown that this heuristic can be also used for other stochastic methods like SGA and ACO for performance improvement.

In short, our proposed AEGA is well-tuned to solved MCDS raised from power graphs, it is 30% faster than SGA, it has the ability to solve large-scale problems, and it achieves better performance than both of SGA and ACO within a specific timeframe.

The rest of this paper is organized as follows: Section 2 review genetic algorithm and its functions and operators. Classification of different kinds of GA algorithms is also provided in this section. Section 3 proposes simple genetic algorithm and its implementation details. Section 4 proposes Adam-Eve like genetic algorithm and details of its implementation, i.e., its functions and procedures. In section 5 first we have formulated MCDS as an optimization problem and then by setting parameters of the proposed Adam-Eve algorithm and defining some heuristics, the proposed method is adjusted to solve MCDS problem, which is derived from WAMS comprehensive design. Simulation results for IEEE 1354 test network, which are obtained by three different meta-heuristics algorithms (i.e. AEGA, SGA, and ACO), will be presented in section 6. This paper will be end with conclusion in section 7.

2. GENETIC ALGORITHM

Holland (29) introduced the genetic algorithm in 1962 and worked on its development with his colleagues during the 1960s and 1970s. In 1975, Holland (29) published a book called “Adaptation in Natural and Artificial Systems.” Since then, extensive research has been conducted on this algorithm, resulting in the introduction of various types of genetic algorithms. The investigation's findings demonstrate the effectiveness of this algorithm in solving diverse problems (12, 13, 29).
The main idea of the genetic algorithm is based on the concepts of natural selection, inheritance, and the possibility of individual change. Its objective is to discover the most effective solution by starting with a set of answers known as the initial population, each represented as a chromosome. To generate the next generation, operators are employed. These operators include selection, mutation, crossover, and replacement, which are applied to each generation in sequence. By utilizing these operators, the genetic algorithm can be executed in a systematic manner (28, 29). An outline of the various steps involved in different genetic algorithms is presented in **Algorithm 1**. Indeed, genetic algorithm utilizes some functions and operators in the aforementioned steps.

2. 1. Genetic Algorithm Functions and Operators

In the subsequent sections, we provide a brief explanation of the roles played by each functions and operators of genetic algorithm.

2. 1. 1. Coding The process of representing solutions to a problem as chromosomes is known as coding. In general, different problems require different coding methods, and the selection of a particular encoding method depends on the nature and characteristic of the problem. Binary encoding and permutation encoding are among the various types of coding available (30). Based on the coding techniques, GAs may be classified as: binary GA, real-valued GA, permutation GA, and tree-Based GA (28, 31).

2. 1. 2. Fitness Function The quality of chromosomes in each generation is determined by the fitness function, which provides a non-negative value for each chromosome based on its ability to solve the main problem. Chromosomes with higher fitness values are more likely to be chosen for the next generation (28, 29).

2. 1. 3. Selection The selection operator in genetic algorithms picks out chromosomes from the present population to act as parents for generating the next generation. Its primary objective is to enhance the quality of the population by selecting worthy chromosomes that will produce superior offspring. The selection process is based on the fitness function, which means that individuals with higher fitness values have a greater chance of being chosen. Various selection techniques are employed, such as roulette wheel, tournament, and random selection (28, 32).

Algorithm 1 Simple Genetic Algorithm

```
Inputs:
    Gen Coding;
    EVALUATE: Fitness Function;
    SELECT: Selection Function;
    REPLACEMENT: Replacement Method;
    SC: Stop Condition;
    $n_{pop}$: Population size;
    $p_c$ and $p_m$: Crossover and Mutation Probabilities.
Output:
    $\mathbf{best}_{\mathrm{sol}}$: Best Solution.

Initialize:
    $[pop] \leftarrow$ Randomly generate $n_{pop}$ number of individuals
    EVALUATE($[pop]$)          ▷ evaluate individuals by fitness function
    gen $\leftarrow 1$                           ▷ first generation

while $\sim$ SC do
    $[pop_c] \leftarrow$ generate individuals by crossover
            ▷ individual selection is based on SELECT and $p_c$
    $[pop_m] \leftarrow$ generate mutated individuals
            ▷ individual selection is based on SELECT and $p_m$
    $[pop_{total}] \leftarrow [pop] \cup [pop_c] \cup [pop_m]$
                        ▷ merge all new and old individuals
    EVALUATE($[pop_{total}]$)
                        ▷ evaluate merged population
    $\mathbf{best}_{\mathrm{sol}} \leftarrow$ best of $[pop_{total}]$
                        ▷ store best solution
    $[pop] \leftarrow$ REPLACEMENT($[pop_{total}]$)
                ▷ select $n_{pop}$ number of merged population
    gen $\leftarrow$ gen + 1
                ▷ preparing for next generation
end while
Return $\mathbf{best}_{\mathrm{sol}}$
```

2. 1. 4. Crossover The function of the crossover operator is to simulate inheritance in successive generations. In order to create the next generation, two chromosomes are chosen from the population as parents, and their chromosomes are merged using the crossover operator, resulting in two offspring. The probability of performing the crossover operator on the entire population is not fixed and is typically set between 60%-90% in SGA. Various methods of crossover exist, such as single point, double-point, and linear crossover, which vary depending on the coding and chromosome type (28, 32, 33).

2. 1. 5. Mutation The mutation operator is used to modify certain genes on a chromosome in order to explore new areas of the search space and discover fresh solutions. This operator has the ability to change the search direction in GA and consequently, this operator can prevent it from getting stuck in local optima. Similar to the crossover, the mutation operator is applied randomly to some chromosomes with a probability (p_m). In SGA, this probability is lower than that of the crossover and typically around 20%. The specific type of mutation employed may vary depending on the coding and chromosome specifications outlined in the algorithm (28-33).

2. 1. 6. Replacement By utilizing crossover and mutation operators, new offspring are generated. During the replacement process, a determination is made regarding which offspring should be included in the population and which parent should be replaced. There are two primary replacement techniques: “Generational Update” and “Steady State Update”. In Generational Update, the number of chromosomes in the population determines the number of offspring produced. This method replaces the previous generation entirely and creates a new population. In Steady State Update, offspring are added to the next-generation population as soon as they are produced. To allow a new chromosome to join the population while maintaining a constant size in SGA, some chromosomes from the existing population must be removed (28, 32).

2. 1. 7. Stop Condition A genetic algorithm is an evolutionary algorithm that operates through iterations and necessitates a stopping point. Various stopping points include the number of generations, constant best fitness for a predetermined number of generations, reaching a predetermined level of fitness, duration of time, and others.

2. 1. 8. Heuristics The genetic algorithm involves creating an initial population randomly, using inheritance (crossover), allowing for generational change (mutation), and natural selection. Depending on the problem at hand, heuristic functions and operators can be utilized in each of these processes to increase efficiency in finding the optimal solution or avoiding local optima. The effectiveness of these heuristics is heavily influenced by factors such as the problem's nature, coding type, fitness function, and other genetic algorithm operators. Examples of commonly used heuristics include generating an initial population within the feasible set that is likely to be close to solutions, elitism replacement, and fascism replacement [23-24].

2. 2. Classification of Genetic Algorithms Previously, it has been noted that genetic algorithms may be classified based on their coding techniques. In general, genetic algorithms can be classified based on various concepts. In terms of implementation, they can be divided into two major categories: sequential and parallel. The sequential ones themselves can be classified as generational, steady-state, and messy (32).

However, genetic algorithms can also be classified based on their evolutions over the generations. In this case, they can be classified into adaptive and non-adaptive genetic algorithms. Adaptive genetic algorithms use feedback from the environment to adjust their parameters and improve performance over time. Non-adaptive genetic algorithms do not change their parameters during the optimization process (27).

3. SIMPLE GENETIC ALGORITHM

The Simple Genetic Algorithm (SGA) is a widely used type of GA that is easy to implement and has an acceptable performance. Several methods have been proposed to implement SGA. The main objective of this study is to implement AEGA, so we start by implementing SGA and then develop it into the Adam-Eve GA, that is to say the SGA presented here has been the basis of the proposed AEGA. Therefore, we need first to describe the details of the implemented SGA.

The flowchart of SGA can be seen in Figure 1. It is assumed that the coding type and fitness function are *a priori* defined. The proposed SGA starts by defining input parameters and then carries out four procedures: `P0`, `P1`, `P2`, and `T`, which generate the initial population, perform crossover, mutation, and replacement, respectively. The next subsections will provide a description of the SGA algorithm.

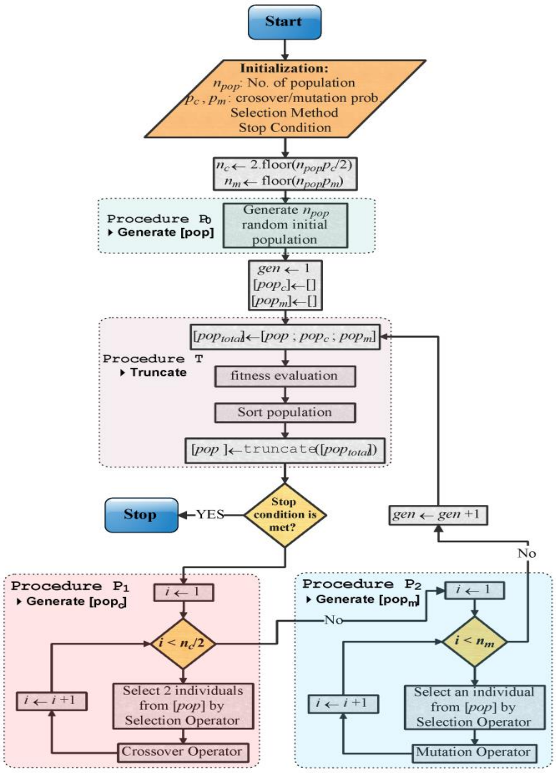


Figure 1. Flowchart of SGA with steady sate update replacement (27)

3. 1. Initialization To begin, the initial values for SGA must be specified. This includes determining the selection function, stop condition, population size (n_{pop}), crossover probability (p_c), and mutation probability (p_m). The proposed SGA utilizes the steady-state update for replacement. Consequently, during the crossover and mutation process, a certain number of offspring, n_c and n_m respectively, are generated as follows.

$$n_c = 2\left\lfloor \frac{n_{pop} \cdot p_c}{2} \right\rfloor, \quad n_m = \left\lfloor n_{pop} \cdot p_m \right\rfloor \tag{1}$$

where, ⌊ . ⌋ is floor function, and due to the creation of two offspring in crossover, the value of n_c is assumed to be even.

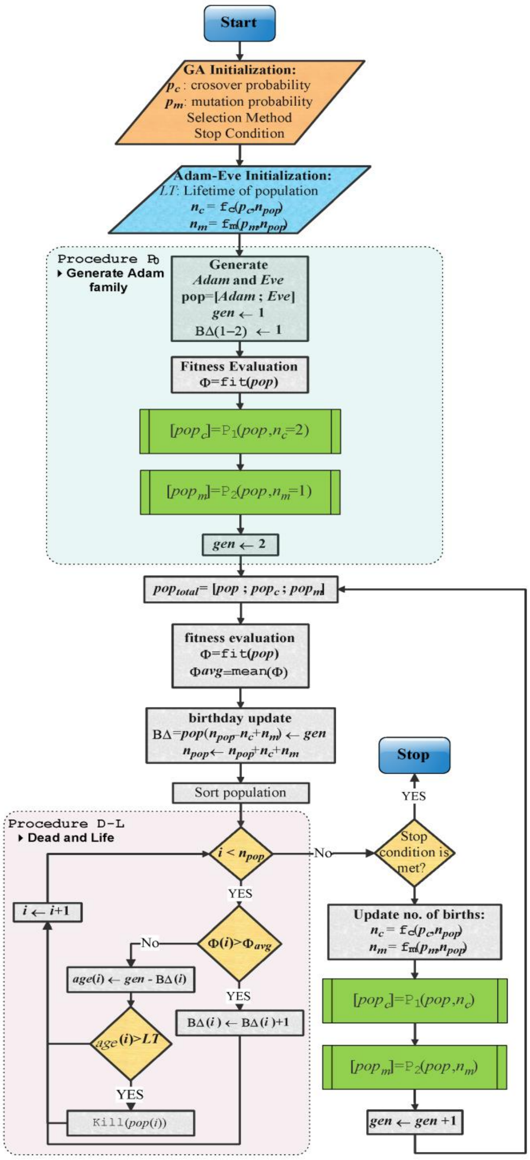


Figure 2. Flowchart of proposed Adam–Eve GA (27)

3. 2. Initial Population and Total Population In the first generation, the n_{pop} number of parents is randomly generated using the `P0()` procedure (population [*pop*]). At this stage, both the population of crossover offspring ([pop_c]) and mutated offspring ([pop_m]) are empty. The total population ([pop_{total}]) is obtained by combining these three populations. It is important to note that in the first generation; [pop_{total}] is equal to [*pop*].

As previously mentioned, one of the heuristics in GA involves generating a portion of the initial population within the feasible set area. This approach increases the likelihood of being closer to potential solutions and greatly impacts convergence speed and performance.

3. 3. Evaluation, Sorting, and Truncation The fitness function is used to evaluate the members of the population created in each generation. These members are then sorted based on their fitness values, from best to worst. In the proposed SGA, the steady state update method is used as the replacement method. From the total population ([pop_{total}]), which consists of $n_{total}=n_{pop}+n_c+n_m$ members, n_{pop} members need to be chosen. This selection is done through the `T()` procedure, known as Truncation. The truncation allows for the implementation of *elitism* and *fascism* heuristics. By defining where n_{pop} members are selected from the sorted population, these heuristics can be applied.

3. 4. Crossover and Mutation Procedures The `P1()` and `P2()` procedures generate populations of crossover and mutated offspring respectively. The offspring are stored in [pop_c] and [pop_m], with n_c and n_m representing their respective numbers. As mentioned earlier, the crossover loop produces two individuals in each run, so n_c is always an even value and the loop iterates $n_c/2$ times.

4. ADAM-EVE LIKE GA AS AN ADAPTIVE GA

The Adam-Eve genetic algorithm is inspired by the biblical story of Adam and Eve, in such a way that the initial population is created by two individuals, referred to as Adam and Eve. These individuals are then used to generate the rest of the population through a process of crossover and mutation. Indeed, the concept of using two initial individuals to generate a population in genetic algorithms has been around since the early days of genetic algorithm research in the 1970s and 1980s. The term "Adam-Eve Genetic Algorithm" may have been coined later as a way to describe this approach.

This algorithm begins with two parents (i.e. Adam and Eve) and as new offspring are born in each generation, the population grows. However, the lifespan of each member of the population must be taken into account as it is natural for individuals to die after their lifetime. This helps to control the rate of population growth and allows for adjustments to be made to keep the population constant. In other words, by adjusting the Adam-Eve parameters effectively and through multiple generations, the number of births is almost balanced by the number of deaths.

After surveying the Adam-Eve algorithm and comparing it to the SGA proposed earlier, it can be inferred that modifying the initial population (`P0()`) and truncation (`T()`) procedures can convert SGA into the Adam-Eve algorithm. Additionally, each generation produces new offspring ($n_c + n_m$) which can be used as parameters for dynamic population control. Upon reviewing the AEGA process, it is evident that this algorithm is adaptive.

The proposed AEGA includes some variables and some procedures, some of which are comparable to the SGA while others are new. The upcoming sections will outline how this algorithm will be put into action, and a visual representation of the algorithm can be seen in Figure 2.

4. 1. Initialization

At the beginning of the AEGA algorithm, initialization is carried out, but unlike SGA, the population size (n_{pop}) is not specified due to the algorithm's characteristics. Apart from SGA, the following elements must also be initialized:

- LT: Life-Time of people;
- $n_c(p_c,n_{pop})$: The function of the number of offspring caused by the crossover;
- $n_m(p_m,n_{pop})$: The function of the number of offspring caused by the mutation.

4. 2. Creation of Adam Family

AEGA differs from SGA in that it starts with Adam and Eve as the first generation, instead of generating a fixed size population randomly. The second generation is then created through two offspring: the one that is created by crossover operator and another that is created by mutation operator. Procedure `P0()` in Figure 2 illustrates this process.

4. 3. Population merging, evaluation, birth registration, and life and death process of Adam Family

First, it is necessary to document the birth details of the descendants in their official record. Then a new generation is created by combining the newly born offspring with their parents. After updating the population and similar to the SGA, it is crucial to assess the fitness of the population. The following are then updated:

- $\phi(i)$: the fitness of i^{th} individual in the population;
- ϕ_{avg}: average of the population fitness;
- $B\Delta(i)$: birth detail of i^{th} descendants in the population.

Next up is the Death and Life process (aka `D-L()` procedure), which involves comparing the fitness levels of individuals to the population average (i.e., ϕ_{avg}). Those with above-average fitness are granted more opportunities for life. After that, each member's age is determined using the following method:

$$\text{age}(i) = \text{gen} - B\Delta(i) \tag{2}$$

where, gen is the current generation number, $B\Delta(i)$ is the generation in which the i^{th} individual was born, and age(i) is the age of the i^{th} individual in the population.

Finally, individuals who have surpassed a certain age threshold (age>LT) are ultimately removed from the population.

5. PROPOSED ADAM–EVE GA FOR WAMS PLANNING

In this section, we will first formulate WAMS comprehensive design as an optimization problem. Then we describe how we have implemented the proposed adaptive genetic algorithm from the previous section to solve this problem.

5. 1. Problem Formulation

The power grids can be expressed by weighted undirected graph. Let $G(V,E,w)$; $w{:}\rightarrow \mathbb{R}^+$ be a weighted undirected graph representing the power grid, where V denotes the electrical buses (i.e. $B_i \in V$) and $ij \in E$ represents the transmission line connecting buses i to j with the length of w_{ij}. The number of buses is considered as the grid size ($|V|=n$), and number of grid connections is denoted by $|E|=M$.

The n-bus power graph can be expressed by n-squared adjacency matrix, where A_w is the weighted one, that is the i^{th} and j^{th} element of this matrix is considered to be w_{ij}, and A is its unweighted version.

The WAMS implemented in $G(V,E,w)$ can be represented by its “measurement infrastructure” and “communication infrastructure”, expressed in Equation 3. The measurement components can be expressed by vector X described in Equation 3.a, which indicates the location of measurements in the system buses. The communication infrastructure of WAMS can be represented by weighted graph $G_{OP}(V_{OP},E_{OP},w)$, as a subgraph of $G(V,E,w)$, where communication nodes, communication links, and length of the links are respectively represented by $h,k \in V_{OP}$; $hk \in E_{OP}$; and w_{hk}. Note that the lengths of a transmission line and its corresponding OPGW link are the same, that is $w=w_{hk}$. Thus, the WAMS can be graphically represented as follows (5):

$$\mathrm{W}(G)=\begin{cases} \mathrm{X}=[\mathrm{x}_i]_{n\times 1} & (3.\mathrm{a}) \\ G_{OP}(V_{OP},E_{OP},w),\ w:E_{OP}\rightarrow \mathrm{R}^+ & (3.\mathrm{b}) \end{cases}$$

$$\texttt{suchthat:}\begin{cases} X\subseteq V_{OP} & (3.\mathrm{c}) \\ V_{OP}\subset V, E_{OP}\subset E & (3.\mathrm{d}) \\ G_{OP}\ \texttt{is connected} & (3.\mathrm{e}) \end{cases}$$

where, n is the size of the network, X denotes PMU location in the system buses, and $G_{OP}(V_{OP},E_{OP})$ is OPGW subgraph.

In general, many applications can be performed by WAMS, such as damping of inter-area oscillation (33), but the main function of PMUs in the power grid has been state observation. The observability analysis indentifies that whether the entire power grid is observable with a set of PMUs or not. The topological observability analysis checks the observability of the entire system by graph theory concepts. The entire system is observable if there is no any zero element in the product of $A^+.X$, where $A^+=A+I_n$ and I_n is identity matrix of size n. It should be noted that "observation" in electrical engineering and "domination" in graph theory are equivalent concepts (5).

The comprehensive design of WAMS aims to place PMUs and OPGWs simultaneously. This can be formulated as the optimization problem as follows (5):

$$\underset{\substack{i\in V(G),\\ hk\in E(G)}}{\text{minimize}} \sum_{i=1}^{n} cp_i.\mathrm{x}_i + \sum\sum_{hk\in E_{OP}} cf.w_{hk}.y_{hk} \qquad (4.\mathrm{a})$$

$$\texttt{subject to:}$$

$$\mathrm{A}^+.\mathrm{X}\succ \hat{1} \qquad (4.\mathrm{b})$$

$$V_{OP}\subset V, E_{OP}\subset E \qquad (4.\mathrm{c})$$

$$X\subseteq V_{OP} \qquad (4.\mathrm{d})$$

$$G_{OP}\ \texttt{is connected} \qquad (4.\mathrm{e})$$

$$\mathrm{x}_i, y_{hk}\in\{0,1\} \qquad (4.\mathrm{f})$$

where, cp_i is the cost of PMU installation at the i^{th} bus, cf is the cost of installation for one kilometer of OPGW, first and second summations are respectively total costs of measurement and communication infrastructures of WAMS. $G_{OP}(V_{OP},E_{OP})$ is a subgraph of power graph G which demonstrates the routers i.e. $\{V_{OP}\}$ and transmission lines equipped by OPGW fiber, i.e. $\{E_{OP}\}$.

The solution of Equation 4 will be minimum connected dominating set of power graph $G(V,E,w)$, where the domination set denotes the PMU locations, and connections specify communication links in the power grids.

It is important to note, as mentioned earlier, that the main issue with MCDS is its multi-objective nature. It aims to minimize both the size of the dominating set and the length of connections between its members. In Equation 4, similar to other work (5, 6), the use of cp_i and cf allows us to transform the MCDS problem into a single-objective problem.

5. 2. Implementation of AEGA for WAMS Design

In this section, we will explain how we have implemented AEGA for the WAMS comprehensive design problem. Additionally, we will describe the heuristics used in the algorithm to improve speed and performance. Note that in adaptive GAs, due to their natures, i.e. adaptation, heuristics are more common than other GAs, especially SGA.

5. 2. 1. Chromosome Coding As explain before, power graph $G(V,E)$ (with the size $n=|V|$, and $M=|E|$ connections) can be expressed by adjacency matrix $A_{n\times n}$. The subgraph $G_{OP}(V_{OP},E_{OP})$ of graph G is also can be expressed by a adjacency matrix A_{OP}. Storing a solution in this way requires n^2 bits, which is not an optimal storage method. As power graph are usually sparse, in the proposed method a M-bit chromosome is suggested that represents connections in the G_{OP} subgraph. Coding and decoding routines enable the conversion of the chromosome to A_{OP} and vice versa.

5. 2. 2. Choice of Adam and Eve In AEGA, the choice of the first two parents (i.e., Adam and Eve) is of great importance. This choice causes AEGA to search the feasible space around these two parents, and their correct selection helps to quickly reach the local optimum. In other words, choosing Adam and Eve intelligently and close to a local optimum (whose its area is approximately known) guarantees finding an optimal solution within only a few generations after Adam and Eve, and this point is valuable in very large-scale problems such as WAMS design.

In the WAMS design, we are looking for the minimum connected dominating set in such a way that a minimum set specifying (aka PMU set) is obtained that are connected to each other, that is, to be connected to each other with the OPGW. On the other hand, we know that a "*Minimum Spanning Tree*" (MST) can be considered as a connected dominating set, because it includes all the nodes and is also connected; but it is not the minimum one. It is also clear that an MCDS does not contain any pendant nodes because it is easy to replace the only node connected to the pendant node and still maintain dominance and reduce the size of the connected subgraph. Based on the given explanations, we go through the following steps to find Adam and Eve:

1. We define the weighted graph $\tilde{G}=(\tilde{V},\tilde{E})$ in such a way that we remove all the pendant nodes from the weighted graph $G(V,E)$.
2. We extract the minimum spanning tree subgraph (which we call M_m) for the weighted graph $\tilde{G}=(\tilde{V},\tilde{E})$; Also, by removing the weight of the

edges of $\tilde{G}=(\tilde{V},\tilde{E})$ and extracting the minimum spanning tree for it, we reach the subgraph M_f.

3. Now we choose M_m as Adam and M_f as Eve.

5. 2. 3. Setting of Probabilities

In general, the crossover operator makes big changes in the chromosomes of the parents, while the mutation operator is a type of fine tuning that makes small changes in the chromosome. The former is a time-consuming process, while the latter is executed faster. In the proposed algorithm, we provided the conditions that the crossover operator is executed with a higher probability at the beginning, and then after a few generations, the probability of crossover decreases and on the contrary, the probability of mutation increases, so that the search continues around the obtained best solution. To achieve this goal, the probability of p_c and p_m is defined as follows. Using these two probabilities, the number of people born by crossover and mutation (n_c and n_m) can be also calculated.

$$p_{c_{new}} = p_c\left(0.85^{\frac{-n_{pop}}{\alpha}}\right) \qquad (5)$$

$$n_c = 2.\lfloor 0.5 n_{pop}.p_{c_{new}} \rfloor$$

where, n_{pop} is number of populations in the current generation. p_c and $p_{c_{new}}$ are respectively probabilities of crossover is the current and next generation. n_c is number of born individuals by crossover operator.

$$p_{m_{new}} = p_m\left(1.25^{\frac{-n_{pop}}{\alpha}}\right) \qquad (6)$$

$$n_m = \lfloor n_{pop}.p_{m_{new}} \rfloor$$

where, p_m and $p_{m_{new}}$ are respectively probabilities of mutation is the current and next generation. n_m is number of born individuals by mutation operator.

5. 2. 4. Elitism

In the genetic algorithm, elitism means giving more chances to elite people in the society. This action increases the speed of reaching the optimum, but at the same time increases the probability of getting stuck in the local optimum. In the proposed AEGA algorithm, it is done by increasing the lifespan of good people, which is fully described in **§4.3**.

5. 3. Innovation of Proposed Algorithm Compared to Previous One

Both algorithms presented in the current research and (28) are Adam and Eve GAs, which are classified in the category of adaptive GAs; with this major difference that the algorithm presented in (28) is designed to solve OPP (i.e., minimum dominating set problem) and the current proposed algorithm to solve OPOP (i.e. minimum connected dominating set problem). As a result of this main difference, there are two basic differences in the algorithm implementation in here and our previous work. The first difference is that the adaptability only occurs in p_c during generations and p_m is a linear function of the population size, while in the current algorithm both p_c and p_m values are adaptive and defined to be exponential functions (5). The second difference is the use of innovative Adam and Eve in the current research, which makes the search process starts from areas closer to the optimal solution. These two mentioned differences cause the proposed algorithm to have the best speed while maintaining efficiency in solving MCDS, which will be shown in the next section with simulations.

6. CASE STUDY AND SIMULATION RESULTS

In order to show the ability of proposed method in WAMS design, IEEE 1354 test case is selected and MCDS problem has been solved by three different meta-heuristic algorithms; AEGA proposed in **§4,** SGA that is proposed in (6) and presented in **§3**, and ACO that is presented in (5). Note that ACO works based on pheromone and visibility of ants, in which evaporation is considered for pheromone. The terms α and β are relative importance power factors between pheromone and visibility function and the term ρ is the evaporation factor and usually is more than %90. For further information about ACO algorithm and its different parameters (α, β, ρ, and Q) we refer the readers to (5).

The parameters of different algorithms are shown in Table 1. We have set ACO parameters based on our previous findings (5). Also, parameter setting for SGA is based on our previous findings (6). Parameter setting in the proposed AEGA is based on Equations 5 and 6, which is discussed before, and the constant values are reported in Table 1. In order to provide same condition for all algorithms, we have used same heuristic in the initial populations of all examined algorithms, that is, we have putted two individuals obtained by the method presented in **§5.2.2** in their initial populations. The selection function is tournament selection in both GAs, the population size of SGA is set to be 70, and number of

TABLE 1. Parameter Settings for Different Algorithms

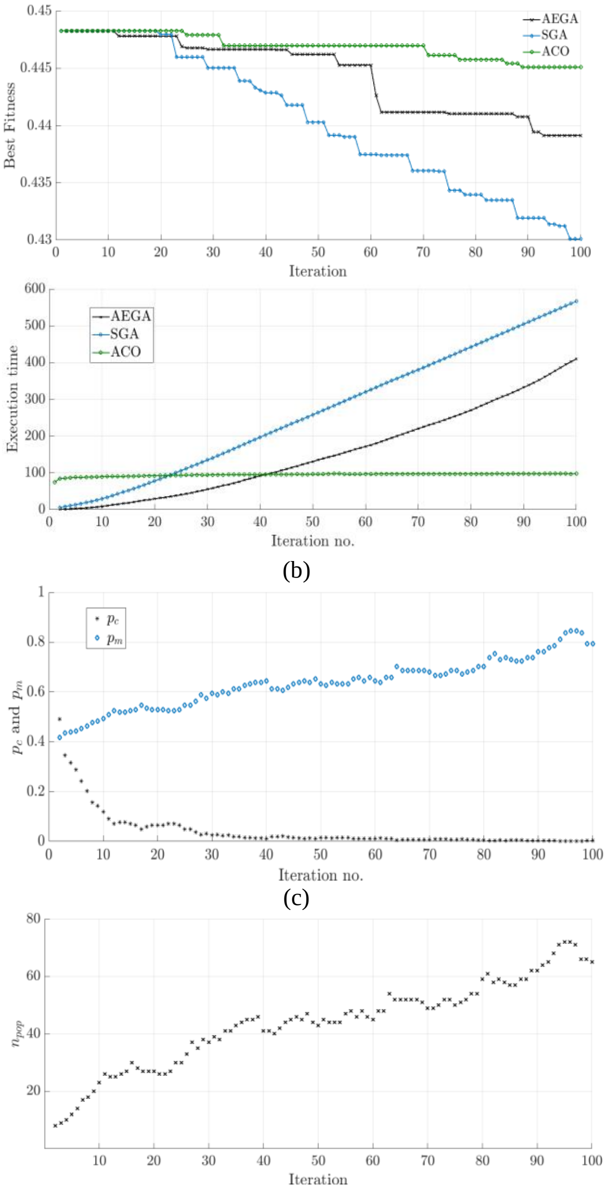
AEGA		SGA		ACO	
p_c	0.6	p_c	0.6	α	1
p_m	0.9	p_m	0.2	β	2
LT	2	n_{pop}	70	ρ	0.9
α	20	-	-	Q	1

ants set to be 20 in ACO. All algorithms have been iterated for 100 times and the evolutions of different parameters in different algorithms are examined during the iterations. Our observations show that AEGA is the most sensitive algorithm to exponential functions of p_c () and p_m (), and finding appropriate functions is another heuristic process. Meanwhile, the least sensitive algorithm is SGA, which has constant values of p_c and p_m, and in most problems, same values are used.

The first parameter is the best solution in each iteration. The best solution is defined as the normalized value of OPGW coverage in percent, i.e., the total length of OPGW fiber for the best solution divided by total length of transmission lines in the power grid. As it can be seen in Figure 3(a), improvement in best solutions occurs first in AEGA (iteration #12), then in SGA (iteration #20), and finally in ACO (iteration #24). This can be considered due to the adaptiveness of AEGA, which adjusts the values of p_c and p_m at the beginning of the search in such a way that the search process is directed faster towards the optimal solution. The rate of improvement in ACO is the lowest among the investigated problems, which is a large-scale network. This is due to the nature of the ant colony algorithm, where the presence of a satisfactory pheromone path (Adam or Eve) results in a low probability of ants moving in new paths, even with high evaporation. Also, the improvement rate in SGA is better than AEGA, which is due to the large number of populations in SGA and spending more time to produce each generation. For this reason, the next parameter that we have evaluated is the production time of each generation.

Figure 3(b) shows the execution times for each iteration in the investigated algorithms (in minutes). It can be seen that the execution time in ACO is almost constant because the ants' tour in a graph with a fixed structure takes a constant time. However, in AEGA, time grows due to population growth, and in SGA, time grows due to the similarity of chromosomes and the process of elitism and fascism. It can also be seen that the execution time for AEGA is always less than that of SGA and about 10% to 30% less than the execution time of SGA. This justifies the low rate of improvement of the best solution in AEGA (shown in Figure 3(a)) because it is able to improve the best solution by spending less time. For example, in the time of 300,000 seconds, AEGA, ACO, and SGA respectively execute 64, 54, and 47 iterations and respectively obtain the best solutions equal to 44.116%, 44.697%, and 44.178%, and this proves that the AEGA algorithm has the better performance at the same execution time.

The next items, which are shown in Figure 3(c), are p_c and p_m probabilities that are confirm the adaptiveness of the proposed AEGA. The decrease of p_c and increase of p_m during different generations can be seen in this figure.



(b)

(c)

(d)

Figure 3. Evolutions of different parameters of Adam-Eve GA, classic GA, and ACO over different iterations: (a) Best fitness for both GAs and ACO; (b) Execution time for different iterations for both GAs and ACO; (c) Variations of p_c and p_m probabilities for different iterations in AEGA; (d) Number of populations over different generations in AEGA

Lastly, the population size in AEGA is shown in Figure 3(d) over different generations. As it can be seen, the population size grows over generations and is fixed at about 70 people in the 100th generation. Actually, this is when the birth and death rates in AEGA become almost equal, and it occurs around the 100th generation due to the adjustments we made in AEGA. That's why we set the population size in SGA to 70 so that the comparison conditions are the same for both GA algorithms.

The results are also shown in detail in Table 2. Considering the importance of changes in the initial generations, 10 iterations are first shown and then with a step of 10.

The closeness of the obtained solution to the optimal solution using the proposed algorithm is another point that should be quantified. Previously, it has been shown that in small and medium-scale power grids, ACO has the best recorded efficiency in finding the solution (5). However, in the current study, it loses its efficiency in large-scale grids due to its slowness. The best solutions for MDS (i.e. OPP reported by Shahraeini (5)) and MCDS (i.e. OPOP) problems are influenced by the structure of power grids and are normally scale-free. To prove this, we have solved the MCDS problem for three small-scale (30, 39, and 57 buses) and three medium-scale (118, 200, and 300) power grids with ACO and reported the results in Table 3. It can be observed that the coverage of edges belonging to MCDS (aka OPGW coverage) in power grids is scale-free and is regularly about 20% to 30% of the total length of power graph edges.

Having the vicinity of the optimal solution in the MCDS problem of the power grids, and considering provided information in **§5.2.2** that the proposed Adam and Eve are naturally shaped as "connected dominating sets" but not the minimum ones, it can be concluded that our heuristic approach in large-scale power grids finds two solutions that are relatively close to the optimal solution (44.828% reported in Table 2). Then, the next step is to implement an algorithm that can guide these solutions to the optimal solution in the shortest time with acceptable efficiency. The results presented in this study indicate that the proposed AEGA algorithm has the best

TABLE 2. Information of each iteration for the proposed AEGA and SGA

gen	Adam-Eve GA						Simple GA		ACO	
	n_{pop}	Best Fitness	Died	p_c	p_m	Execution Time (sec)	Best Fitness	Execution Time (sec)	Best Fitness	Execution Time (sec)
1	-	-	-	-	-	-	0.44828	181.4462	0.44828	4480.5533
2	8	0.44828	0	0.49127	0.41699	30.2831	0.44828	336.2431	0.44828	5049.0588
3	9	0.44828	2	0.34479	0.43469	70.234	0.44828	500.7388	0.44828	5124.0784
4	10	0.44828	2	0.31558	0.43924	115.8274	0.44828	635.3813	0.44828	5195.838
5	12	0.44828	1	0.28884	0.44383	163.4357	0.44828	797.9163	0.44828	5251.6817
6	14	0.44828	1	0.24198	0.45315	218.0468	0.44828	957.5509	0.44828	5247.2689
7	17	0.44828	0	0.20272	0.46268	278.5821	0.44828	1156.554	0.44828	5290.2011
8	18	0.44828	2	0.15544	0.47733	348.7935	0.44828	1334.2016	0.44828	5293.0065
9	20	0.44828	1	0.14227	0.48232	426.1163	0.44828	1548.3312	0.44828	5322.3575
10	23	0.44828	0	0.11919	0.49246	519.0996	0.44828	1767.5569	0.44828	5358.0243
20	27	0.44781	4	0.06414	0.52964	1765.4013	0.44796	4677.2578	0.44828	5531.1844
30	37	0.44666	4	0.024224	0.59381	3312.1561	0.44504	8089.361	0.44791	5648.0553
40	41	0.44666	8	0.011932	0.64531	5501.8556	0.44286	11812.5558	0.44697	5722.874
50	43	0.44622	4	0.014243	0.63203	7852.3392	0.44029	15479.6895	0.44697	5768.0213
60	45	0.44528	4	0.011932	0.64531	10300.7435	0.43747	19196.2619	0.44697	5806.1309
70	49	0.44116	5	0.0076646	0.67975	13217.683	0.43606	22819.7922	0.44697	5823.0418
80	59	0.44102	1	0.005877	0.70128	16200.4046	0.43396	26519.5783	0.44576	5824.9127
90	62	0.44076	3	0.0028947	0.76211	19920.6118	0.43192	30252.2871	0.44509	5841.1705
100	65	0.43914	6	0.0020316	0.79447	24565.5268	0.4289	33982.6806	0.44509	5846.1951

TABLE 3. Best Solution of MCDS for six small/medium grids

Test Case	OPGW Cov. (%)
30-bus	21.359
39-bus	37.285
57-bus	29.505
118-bus	23.357
200-bus	24.813
300-bus	15.672

speed to reach the solution while maintaining acceptable efficiency. However, the best solution is obtained with SGA at the cost of high time consumption. On the other hand, ACO lacks the necessary efficiency even with very high time consumption.

7. CONCLUSION

The WAMS comprehensive design aims to simultaneously place measurements and their required communication links. This can be formulated as an optimization problem and the output will be minimum connected dominating set, which has been a well-known graph problem and due to its NP-hardness, evolutionary algorithms are preferred for solving.

This study proposes an adaptive genetic algorithm to solve MCDS problem, which is knows as Adam-Eve like genetic algorithm. Designing an adaptive genetic algorithm is inherently more challenging than the simple genetic algorithm, and its success relies on the specific characteristics of the problem being studied, which is well done for MCDS problem in the current study.

We have designed the algorithm in such a way that it starts with two parents close to the optimum, and in the initial generations, they are more likely to crossover, and with the passage of generations, this probability decreases and, conversely, the probability of mutation increases in order to reach the optimum by fine tuning.

The simulation results for IEEE 1354 bus test network indicate that our AEGA has been up to 30% faster than SGA, while the efficiency of the algorithm is acceptable. This speed improvement can be used in other domain like wireless sensor networks, where the speed is more important than the efficiency. Also we have observed that ACO has not the capability of solving MCDS problems in the large-scale power graphs.

Our recent researches have demonstrated that the adaptive genetic algorithm is capable of enhancing the resolution of complex network problems, as evidenced by the speed and performance enhancement in simultaneous placement of WAMS components (i.e., PMUs and OPGWs) in the current study and the performance enhancement in PMU placement from our previous research.

Further investigation should be done to the appropriate selection of Adam and Eve, as well as the adaptability of the two functions of crossover and mutation and the process of death of individuals, in order to increase the efficiency of the proposed algorithm. The current study has been only focus on the WAMS comprehensive design, while adaptive GAs can be also used in the sequential methods that have previously proposed for WAMS design.

8. ACKNOWLEDGMENT

This paper is an extended version of the conference paper which was selected as one of the best paper award finalists in *2022 8th Iranian Conference on Signal Processing and Intelligent Systems (ICSPIS)*, Mazandaran, Iran.

9. REFERENCES

1. Maheswari M, Suthanthira Vanitha N, Loganathan N. Wide-area measurement systems and phasor measurement units. Wide Area power systems stability, protection, and security. 2021:105-26. 10.1007/978-3-030-54275-7
2. Shahraeini M, Javidi MH, Haq Z. Wide area measurement systems. Advanced topics in measurements. 2012:303-22. 10.5772/35466
3. Maji TK, Acharjee P. Multiple solutions of optimal PMU placement using exponential binary PSO algorithm for smart grid applications. IEEE transactions on industry applications. 2017;53(3):2550-9. 10.1109/TIA.2017.2666091
4. Shahraeini M, Javidi MH, Ghazizadeh MS. Communication infrastructure planning for wide area measurement systems in power systems. International Journal of Communication Networks and Distributed Systems. 2013;10(4):319-34. 10.1504/IJCNDS.2013.054229
5. Shahraeini M, Kohsari G, Javidi MH, editors. Comparison of Meta-Heuristic Algorithms for Solving Dominating Set Problems in WAMS Design. 2022 8th Iranian Conference on Signal Processing and Intelligent Systems (ICSPIS); 2022: IEEE. 10.1109/ICSPIS56952.2022.10044037
6. Shahraeini M, Ghazizadeh MS, Javidi MH. Co-optimal placement of measurement devices and their related communication infrastructure in wide area measurement systems. IEEE Transactions on Smart Grid. 2012;3(2):684-91. 10.1109/TSG.2011.2178080
7. Kvasov DE, Mukhametzhanov MS. Metaheuristic vs. deterministic global optimization algorithms: The univariate case. Applied Mathematics and Computation. 2018;318:245-59. 10.1016/j.amc.2017.05.014
8. Colaço MJ, Dulikravich GS. A survey of basic deterministic, heuristic and hybrid methods for single-objective optimization and response surface generation. Thermal measurements and inverse techniques. 2011;1:355-405.
9. Khalkhali A, Narimanzadeh N, Khakshournia S, Amiri S. Optimal design of sandwich panels using multi-objective genetic algorithm and finite element method. International Journal of Engineering, Transactions C: Aspects 2014;27(3):395-402. 10.5829/idosi.ije.2014.27.03c.06
10. Arab R, Ghaderi S, Tavakkoli-Moghaddam R. Solving a new multi-objective inventory-routing problem by a non-dominated sorting genetic algorithm. International Journal of Engineering, Transactions A: Basics. 2018;31(4):588-96. 10.5829/ije.2018.31.04a.10
11. Khodadadi M, Kazemi A, Gorji M. Optimal thermodynamic design of turbofan engines using multi-objective genetic algorithm. International Journal of Engineering, Transactions C: Aspects. 2014;27(6):961-70. 10.5829/idosi.ije.2014.27.06c.15
12. Sivanandam S, Deepa S, Sivanandam S, Deepa S. Classification of genetic algorithm. Introduction to Genetic Algorithms. 2008:105-29. 10.1007/978-3-540-73190-0_5

13. Feyzbakhsh SA, Matsui M. Adam–Eve-like genetic algorithm: a methodology for optimal design of a simple flexible assembly system. Computers & Industrial Engineering. 1999;36(2):233-58. 10.1016/S0360-8352(99)00131-X
14. Li B, Zhang X, Cai S, Lin J, Wang Y, Blum C, editors. Nucds: An efficient local search algorithm for minimum connected dominating set. Proceedings of the Twenty-Ninth International Conference on International Joint Conferences on Artificial Intelligence; 2021. 10.24963/ijcai.2020/209
15. Mohammadi MB, Hooshmand R-A, Fesharaki FH. A new approach for optimal placement of PMUs and their required communication infrastructure in order to minimize the cost of the WAMS. IEEE Transactions on smart grid. 2015;7(1):84-93. 10.1109/TSG.2015.2404855
16. Ghasemkhani A, Monsef H, Rahimi-Kian A, Anvari-Moghaddam A. Optimal design of a wide area measurement system for improvement of power network monitoring using a dynamic multiobjective shortest path algorithm. IEEE Systems Journal. 2015;11(4):2303-14. 10.1109/JSYST.2015.2469742
17. Dubey R, Popov M, Muro JdJC. Cost effective wide area measurement systems for smart power network. IEEE Power and Energy Technology Systems Journal. 2018;5(3):85-93. 10.1109/JPETS.2018.2835646
18. Appasani B, Mohanta DK. Co-optimal placement of PMUs and their communication infrastructure for minimization of propagation delay in the WAMS. IEEE Transactions on Industrial Informatics. 2018;14(5):2120-32. 10.1109/TII.2018.2799659
19. Fan N, Watson J-P, editors. Solving the connected dominating set problem and power dominating set problem by integer programming. Combinatorial Optimization and Applications: 6th International Conference, COCOA 2012, Banff, AB, Canada, August 5-9, 2012 Proceedings 6; 2012: Springer. 10.1007/978-3-642-31770-5_33
20. Kucuktezcan CF, Genc VI. A new dynamic security enhancement method via genetic algorithms integrated with neural network based tools. Electric Power Systems Research. 2012;83(1):1-8. 10.1016/j.epsr.2011.09.004
21. Kaya M. The effects of two new crossover operators on genetic algorithm performance. Applied soft computing. 2011;11(1):881-90.
22. Hassanat A, Almohammadi K, Alkafaween Ea, Abunawas E, Hammouri A, Prasath VS. Choosing mutation and crossover ratios for genetic algorithms—a review with a new dynamic approach. Information. 2019;10(12):390. 10.3390/info10120390
23. Michalewicz Z. Heuristic methods for evolutionary computation techniques. Journal of Heuristics. 1996;1:177-206. 10.1007/BF00127077
24. Tovey CA. Nature-inspired heuristics: Overview and critique. Recent advances in optimization and modeling of contemporary problems. 2018:158-92. 10.1287/educ.2018.0187
25. Putnins M, Androulakis IP. Self-selection of evolutionary strategies: adaptive versus non-adaptive forces. Heliyon. 2021;7(5). 10.1016/j.heliyon.2021.e06997
26. Alexandrov IA, Kirichek AV, Kuklin VZ, Chervyakov LM. Development of an Algorithm for Multicriteria Optimization of Deep Learning Neural Networks. HighTech and Innovation Journal. 2023;4(1):157-73. 10.28991/HIJ-2023-04-01-011
27. lal Rajora G, Sanz-Bobi MA, Domingo CM. Application of Machine Learning Methods for Asset Management on Power Distribution Networks. Emerging Science Journal. 2022;6(4):905-20. 10.28991/ESJ-2022-06-04-017
28. Shahraeini M, Soltanifar R, editors. Performance Comparison between Simple and Adam—Eve-Like Genetic Algorithms in Optimal PMU Placement Problem. 2022 8th Iranian Conference on Signal Processing and Intelligent Systems (ICSPIS); 2022: IEEE. 10.1109/ICSPIS56952.2022.10043948
29. Holland J. Adaptation in neural and artificial system. Ann Arbor, Univeristy of Michigan Press. 1975.
30. Hasibuan A, Siregar WV, Isa M, Warman E, Finata R, Mursalin M. The Use of Regression Method on Simple E for Estimating Electrical Energy Consumption. HighTech and Innovation Journal. 2022;3(3):306-18. 10.28991/HIJ-SP2022-03-06
31. Goldberg D. Genetic Algorithms in Search, Optimization and Machine Learning,. Addition-Westly. Reading MA. 1989.
32. Deepa S, Sivanandam S. Introduction to genetic algorithms: Springer; 2010.
33. Asghari R, Mozafari S, Amraee T. Delay-scheduled controllers for inter-area oscillations considering time delays. International Journal of Engineering, Transactions B: Applications. 2018;31(11):1852-61. 10.5829/ije.2018.31.11b.08

Persian Abstract

چکیده

سیستم اندازه‌گیری ناحیه گسترده (WAMS) از دو زیرساخت مختلف اندازه‌گیری و مخابراتی تشکیل شده است که وظیفه اندازه‌گیری داده‌های شبکه در ناحیه گسترده و ارسال و پردازش آنها در مراکز کنترل را بر عهده دارد. طراحی WAMS می‌تواند شامل طراحی هر کدام از زیرساخت‌های آن باشد و یا هر دو زیرساخت را همزمان هدف قرار دهد که به طراحی جامع WAMS مشهور است. طراحی جامع WAMS به معنی جانمائی همزمان ادوات اندازه‌گیری و ارتباطات مورد نیاز آن است که در تئوری گراف‌ها به عنوان مجموعه متصل غالب کمینه (MCDS) شناخته می‌شود و در قالب یک مساله بهینه‌سازی فرمول‌بندی می‌شود. حل چنین مساله بهینه‌سازی پیچیده‌ای غالبا با الگوریتم‌های جمعیت‌گرا مانند الگوریتم ژنتیک انجام می‌شود و سرعت و کارائی یافتن پاسخ همواره یک چالش بوده است. این پژوهش یک الگوریتم ژنتیک تطبیقی موسوم به الگوریتم آدم و حوا پیشنهاد می‌کند که قابلیت حل مساله MCDS که از طراحی جامع WAMS ناشی شده را دارد. از طریق نتایج شبیه‌سازی برای شبکه IEEE ۱۳۵٤ باس نشان می‌دهیم که الگوریتم پیشنهادی به خوبی برای حل MCDS مرتبط با گراف‌های قدرت طراحی و تنظیم شده است. این الگوریتم ۳۰٪ سریعتر از الگوریتم ژنتیک ساده است، مسائل مقیاس-بزرگ را به طور موثر حل می‌کند، و در یک بازه زمانی معین از الگوریتم ژنتیک ساده و الگوریتم کلنی مورچه‌ها بهتر عمل می‌کند.

IJE TRANSACTIONS B: Applications Vol. 37 No. 02, (February 2024) 312-322

International Journal of Engineering

Journal Homepage: www.ije.ir

Implementation of Chatbot that Predicts an Illness Dynamically using Machine Learning Techniques

S. Shedthi B.[a], V. Shetty*[b], R. Chadaga[a], R. Bhat[a], B. Preethi[a], P. Kini K.[a]

[a] *Nitte (Deemed to be University), NMAM Institute of Technology (NMAMIT), Department of Computer Science and Engineering, Nitte, India*
[b] *Nitte (Deemed to be University), NMAM Institute of Technology (NMAMIT), Department of Mechanical Engineering, Nitte, India*

PAPER INFO

Paper history:
Received 11 August 2023
Received in revised form 11 September 2023
Accepted 12 September 2023

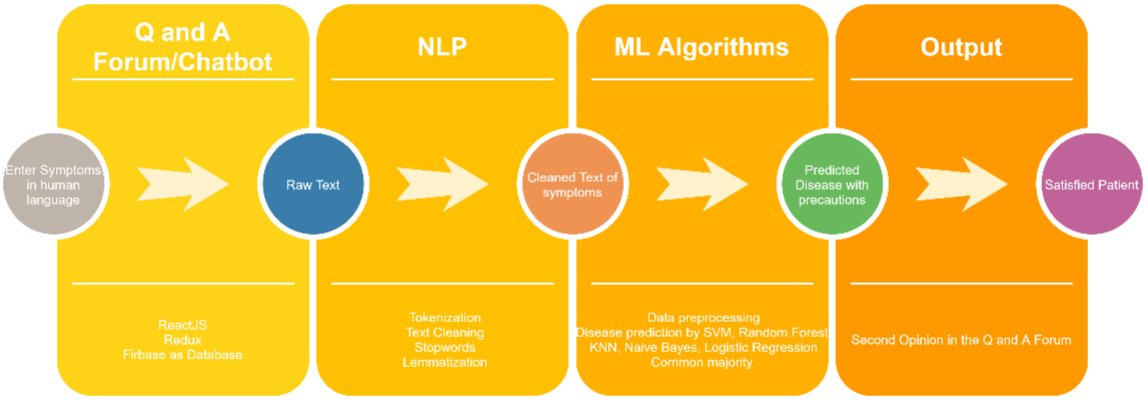
Keywords:
Artificial Intelligence
Machine Learning
Natural Language Processing
Healthcare
Chatbot

ABSTRACT

Timely access to healthcare is crucial in order to maintain a high standard of living. However, obtaining medical consultations can be difficult, especially for those living in remote areas or during a pandemic when face-to-face consultations are not always possible. The ability to accurately diagnose diseases is essential for effective treatment, and recent technological advancements offer a potential solution. Machine learning (ML) and Natural language processing (NLP) enables computer programs to understand human language and extract desired features from responses, allowing for human-like interaction with users. By leveraging these technologies, healthcare professionals can potentially provide more accessible and efficient medical consultations to individuals, regardless of their location. The concept is to establish an online platform where users can ask medical-related queries and receive responses from both medical professionals and fellow users. The platform would feature a Medical Chatbot, which employs advanced ML techniques to analyze user-provided symptoms and provide initial disease diagnosis and related information prior to consulting with a doctor. This disease prediction chatbot interacts dynamically with the users to enter the symptoms of the diseases and based on syntactic and semantic similarity response is given. In this work the threshold of similarity score is kept of 0.7. K-Nearest neighbors, Random forest, Support vector machine, Naive bayes and Logistic regression algorithms are used for prediction of disease based on symptoms which are faced by users. The syntactic similarity, fuzzy string matching and semantic similarity using all-MiniLM-L6-v2 model is used to improve the efficiency of the result.

***doi**: 10.5829/ije.2024.37.02b.08*

Graphical Abstract



---

*Corresponding Author Email: *vidyasagar.shetty@gmail.com* (V. Shetty)

Please cite this article as: Shedthi B. S, Shetty V, Chadaga R, Bhat R, Preethi B, Kini K P. Implementation of Chatbot that Predicts an Illness Dynamically using Machine Learning Techniques, International Journal of Engineering, Transactions B: Applications. 2024;37(02):312-22.

1. INTRODUCTION

Machine learning is an area of artificial intelligence (AI) with a concept that a computer program can learn and adapt to new data without human intervention (1). When a machine learning system receives new data, it guesses the outcome using the prediction models it has built using prior data. The amount of data used determines how well the output is anticipated, as a larger data set makes it easier to create a model that predicts the outcome more precisely. Natural Language Processing (NLP) is a subfield of AI which uses ML algorithms to train the computers to understand, interpret human language and derive meaningful insight from it (2). NLP uses computer algorithms with machine learning to process and analyze textual data which helps to analyze the sentiments (3-5). NLP combines techniques from linguistics, computer science, and machine learning to process and analyze natural language data, such as text or speech has numerous applications in areas such as information retrieval, sentiment analysis, Chatbots, and language translation. By enabling computers to process and understand natural language, NLP is bridging the gap between humans and machines and opening up new channels for connection and communication between the people (6). Chatbots are an intelligent system being developed using AI and NLP algorithms (7).

A chatbot can also be known as digital assistants that understand human capabilities and it interpret the user intent, process their requests, and give prompt relevant answers. Chatbots have become increasingly popular as useful tools for companies and institutions. Many companies recognize the potential for using Chatbots in healthcare to support patients. Obtaining consultations with doctors for every health issue can be challenging. The concept is to develop a medical chatbot using Artificial Intelligence, which can diagnose diseases and provide essential information about them before consulting a doctor. This approach aims to decrease healthcare expenses and enhance access to medical knowledge through the use of a medical chatbot (8). Chatbots are computer programs that interact with users using natural language. They store data in a database to identify keywords in sentences and provide appropriate responses to user queries.

In 1950, Alan Turing introduced the concept of Chatbots by posing the question, "Can machines think?" (9). The first known chatbot, ELIZA, was created to act as a psychotherapist and utilized pattern matching and template-based responses to engage in question-based conversations (10). Another well-known chatbot, ALICE, was developed later and employed a pattern-matching technique to retrieve example sentences from output templates and avoid inappropriate responses (11). This concept was then expanded to various fields using machine learning techniques. Among the most popular applications of Chatbots in healthcare are diagnostics, patient support and health promotion (12).

A medical Chatbot is a conversational AI powered solution specifically designed to make healthcare much more interactive and proactive. Objective of the proposed work is to develop an effective question-and-answer forum where users can discuss the queries on the medical field and also develop a medical Chatbot that predicts an illness dynamically based on user symptoms using NLP and machine learning techniques.

Section 2 provides an overview of previous research conducted on disease diagnosis using machine learning and also about Chatbots. In Section 3, we discuss the fundamental principles of Chatbot and the construction of our proposed work. Section 4 is dedicated to the implementation and results of the Question and Answer Forum and Chatbot model. Finally, in section 5, we present a summary of our findings and offer suggestions for future work.

2. RELATED WORK

In this literature survey section, a detailed summary of the different kinds of research that have been conducted that are related to disease diagnosis using machine learning and Chatbots are considered.

Problem-Based Learning (PBL) is a key point for learning activities (13) and technology plays an important role in the learning process (14). Computer science and health care field is correlated and lots of research is going on with combination of these two fields (15-17). Ahsan et al. (18), this paper is a comprehensive review of various algorithms in Machine Learning for popular diseases based on symptoms. There is also a great raise in the number of research papers published from 2012 to 2021 on different diseases. This review also mentions the possible data related problems like data scarcity, noisy data and manipulation of training data making us aware to be careful of the data set selection, as well as disease diagnosis related challenges like misclassification and confusion that algorithms predict. This paper is very insightful as it provided history of many diseases related classifications and where algorithms could be inaccurate. Shaji et al. (19) investigated on heart diseases which are getting extremely common and there are a variety of attributes such as age, sex, blood pressure, and the levels of urea, sodium and potassium will help in predicting, if a person is suffering from a heart disease. The factors like stress level or if any person is a smoker or not plays a major role in determining the disease. These attributes were taken as data and used in classifying algorithms like SVM, RF, KNN and ANN. This paper though very specific to detection of heart diseases and its symptoms, gives a great direction on how the process of data cleaning is done and compares the accuracy of algorithms and is very helpful for researchers starting out in the medical AI and ML field.

Keniya et al. (20) used various KNN and gave the most accurate prediction with accuracy of 93%. It used attributes like symptoms, age and gender to determine a disease out of 230 possible outcomes. The aim of this work was that sometimes doctors are unavailable and diseases like Ebola are deadly and contagious so we can avoid human contact as much as possible which matches our aim too. This project lacks the User Interface which usually makes access easier to healthcare professionals.

Ferjani et al. (21) proposed supervised ML algorithms for disease diagnosis and aiding medical experts in the early identification of high-risk diseases. The aim of this work was to investigate and analyze already existing algorithms and focuses on the accuracy of commonly used algorithms and find out which algorithm is more precise for which of the diseases. Magoulas et al. (22) insisted that most disease prediction systems are application driven. It can face numerous issues like overfitting and improper scaling. The visualization of the learnt problem will be tough. We also need to deal with large datasets which may have high dimensional input features. Also, there are ethical aspects in the medical field and human intervention is necessary. This work indicates AI and ML in healthcare is no new topic. This was way ahead of the times it was published, almost all statements are true even now. Though there are various python libraries to help us with the visualization. Jain et al. (22) stated that people do not go to the doctor for harmless diseases like cold, sneezing, fever which may be a symptom of a more harmful disease. The person must choose symptoms from groups of it which range from mild, medium, and severe symptoms. It uses DT and deep neural networks to classify the disease as well as compares the two algorithms used. Researchers (21-23) are given an in-depth detail of AI and ML in healthcare, how to deal with data. The algorithms to try on our data set and driving the force that prompted the backend of our project. Pingale et al. (24) used Naïve Bayes, KNN and Logistic Regression to diagnose a disease and the dataset is from real-life hospital data which is in structural format. This work highlights two points that our work also stresses in the sections mentioned below, that if a user enters unrelated or random symptoms or if there are a smaller number of symptoms, there are high chances of obtaining less accurate results.

Bharti et al. (25) proposes a multilingual conversational bot which intends to give free primary healthcare and education. It is implemented on google cloud platform and uses NLP to implement the system. The paper also talks about voice-based user interface as well as a chat-based interface. It's based on DialogFlow and uses only 255 intents which may be a limitation if put into production. Though this work has a great interface including a message based and voice-based AI doctor; this is not very helpful to healthcare professionals who need to deal with more complicated diseases. It focuses on the use where possible patients need advice. A review paper by Tjiptomongsoguno et al. (26) which makes a logical table comparing the Chatbots based on the methodology, algorithms and the datasets used. It also highlights the various pros and cons found in each Chatbot. This work highlights the necessity of medical websites in today's digital world and was the root cause that a question-and-answer forum/website exists in our work. Reshma et al. (27) uses SVM to predict health status and it also uses the google API to convert speech to text and text to voice. The input is sent to a Chatbot, which responds with relevant information and displays it on the stand-alone app. Prayitno et al. (28) proposed a Chatbot for frequently asked questions in a company which uses NLP to successfully diagnose the user's illness with 87 percent accuracy.

Ayanouz et al. (29) performed a detailed survey and examined many publications which are related to Chatbots using the AI concepts needed to build an intelligent conversational agent based on Deep Learning Models. One of the Chatbots is Casper which helps Insomniacs pass the night. Another is One Remission Chatbot which has the goal of helping those involved in the fight against cancer. This paper also mentions that selecting the right engine for NLP is the most important step of creating a Chatbot. Programmers also must decide whether they want structured or unstructured conversations. They also tell the basics of NLP and finally propose a general architecture of a smart Chatbot.

Caldarini et al. (30) mainly classifies Chatbots into rule-based Chatbots and Artificial intelligent Chatbots. Though rule-based models were easy to design and implement, they had limited capabilities. AI Models are based on ML Algorithms, and they allow them to learn from an existing database of human conversations. It also talks about different datasets used like Open Subtitles, Cornell, Daily Dialog. A variety of evaluation metrics used are F1 score, Perplexity and Bilingual evaluation understudy (BLEU). Kumar et al. (31) concentrates on NLP and neural network and summarizes the most efficient implementation techniques that have been carried out in the previous years. It then proposes a methodology to develop a state-of-the-art Chatbot application that can be personalized easily according to customer needs. It is implemented using tools such as Dialog Flow, Tensor Flow, Android Studio and Firebase. Chaudhary et al. (32), in this paper presented the pros and cons along with comparisons of various available medical Chatbots. A multilingual Chatbot called 'HEALTHBOT' was also proposed which will interact with patients in English and Marathi and will note down their symptoms and pathological test reports and will also prescribe the patients with further medical tests and suggest them basic medications, diets, and lifestyle changes. Also is highly useful for the medical practitioners as personal assistant and not only that it would save valuable time for both the parties and reduce the treatment time.

Vasileiou and Maglogiannis (33) developed a health Chatbot predict and diagnosis and limitation of this work is it is used for only two diseases i.e., Covid-19 and heart problem. In future work we can extend this work with more diseases. Shaikh et al. (34) presented an adolescent oriented intelligent conversational chatting system called “HappySoul”, which acts as a virtual friend who can assist to encourage, understand, comfort, and guide stressful adolescents to pour out their bad and negative feelings, thereby releasing the stress. Chatbot will allow a user to simply ask questions in the same way that they would address a human. The technologies used were NLP, RNN and client server architecture with the help of Android GUI. Above works made us realize that using NLP even though more difficult than DialogFlow and TensorFlow (They need us to define intents) gives us more independence and flexibility in the design of the Chatbot. PhaniRaghavaa and Kumarb (35) aimed to create an artificial conversation entity for healthcare treatment using python. Two algorithms, i.e., fuzzy support vector machine algorithms are compared with the Decision tree algorithm. The results indicate that the suggested fuzzy support vector machine algorithm will outperform the current approach in terms of output.

Eslam, et al. (36) presented a smart Chatbot system that can communicate with people, and it gives answers about the COVID-19. The model used the pretrained Google BERT language model. This technique employs the BERT Transformer to categorize text input into various categories based on the meaning of the words. The second stage involves using the BERT model and choosing the query domain for the replies. Their proposed system is trained and tested on Stanford University’s SQuADV2.0, a well-known question-answering dataset. Kumar et al. (37) have used the classification algorithms such as Naive Bayes, Random Forest, Logistic Regression, and KNN to predict the disease based on the symptoms inputted by the patient. These algorithms have different accuracies and are chosen based on the specific disease being predicted. Moreover, the system can also determine whether a patient is suffering from a specific disease or not by predicting “True” or “False”. Once the disease is predicted, the system can recommend the patient which type of doctor to consult for the specific disease. The entire model has been implemented using Django and connected to the Django server. Tamizharasi et al. (38) proposed to make accurate predictions and increase the model's effectiveness. In their study, they addressed the use of a support vector machine learning method in a Chatbot for healthcare. The conversational style is achieved using natural language processing. In addition to receiving free or low-cost services, this helps people spend less time.

In recent work lots of improvement done related to healthcare digitalization in that Chatbot is a prominent one (39). There is no doubting the extent to which the use of AI, including Chatbots, will continue to grow in public health (40). In an ideal scenario, research should play a crucial role in shaping the progress of digital innovations in public health. This involves assessing the effects of technologies like Chatbots on health interventions, gaining understanding of user experiences, and giving utmost importance to the safety and welfare of users.

From the above literature survey observed that NLP and machine learning technology can help to give solution to most healthcare problems. But in most of the works they have used dialogflow and tensor flow for Chatbot creation and the final test result is derived through one best training algorithm. Further work can be extended on developing a robust experimental design that can effectively demonstrate the efficacy and efficiency of the Chatbot. Additionally, it is necessary to create a user-friendly Question and Answer Forum to facilitate better interaction.

3. PROPOSED SYSTEM

There sometimes arises a situation where a person is feeling ill but cannot access a doctor in a clinic or hospital because of odd timings or a person is in a remote area. Instead of panicking and self-diagnosing randomly it is better to have knowledge of possible diseases a person is suffering from and take precautions for the same.

The aim of this research work is to design a platform where users can ask health related queries is important to have a wider view on the conditions. Search, posting questions, answering questions if relevant are the additional features which can escalate its use. A medical Chatbot can also help in predicting the disease given the symptoms. Our project aims at developing a website where users can have an interaction with fellow users on health-related aspects, and an integrated medical Chatbot created using machine learning and natural language processing techniques. People can interact with the Chatbot just like they do with another human and through a series of queries; Chatbot will identify the symptoms of the user and thereby, predict the disease using a machine learning algorithm. The proposed system hence reduces hospital wait times, consultation time, hospital readmission time in case patients need to be connected to the correct healthcare professionals and helps them understand their conditions and possible treatment option without even visiting a doctor. Front end design of the proposed system is shown in Figure 1. Redux and ReactJS used for Chatbot front end.

Redux and ReactJS are both popular tools for managing the state of a web application, but they approach the problem in different ways. ReactJS has its own built-in state management system, known as local state. Local state is managed within a component and is not shared with other components. Redux, on the other hand, is a standalone state management library that can be used with React or other frameworks. Redux provides a central store that holds the entire state of the application, which can be accessed and modified by any component that subscribes to the store. To update the state, you dispatch an action to the store, which triggers a series of reducers that modify the state in a predictable and consistent way. React is one of the most popular frontend JavaScript Library. Even though there are other popular frameworks like Angular JS, Vue JS, Ember JS and many more. There were certain characteristics of React and Redux giving them the upper hand and hence they were chosen as our framework. React JS uses a virtual DOM (Document Object Model) that updates only parts of the DOM that were changed whereas Angular JS uses real DOM which will be slower while rendering large amounts of data. Thus React and Redux have better performance than Angular JS. React and Redux also have more flexibility than Vue JS. React JS has better support for mobile development. ReactJS and Redux are scalable, a small application can be gradually converted to a larger application while Ember JS is for already scaled systems.

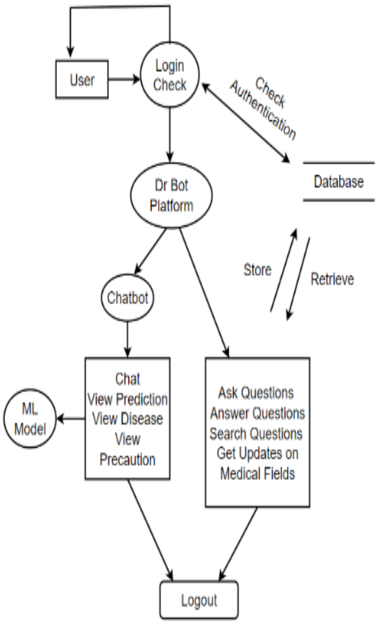


Figure 1. Chatbot front end using Redux and ReactJS

Chatbot Backend done using NLP and steps of that is shown in Figure 2. The purpose of using NLP for the chat bot is that NLP ensures that the Chatbot can understand the vocabulary, sentiments and meaning that users use when naturally conversing. There was a possibility of that the project could be done without the use of NLP by presenting the symptoms as options to select by the users. This would remove the ease of use and make it harder for the users to choose their symptoms as people may not even have certain symptoms. For example, the symptom nausea means having discomfort and feeling the need to vomit. A person who does not know “nausea” may input “discomfort” and “vomiting” and NLP can make semantic relation between them to choose “nausea” as the symptom. The SpaCy Library, which is a free, open source library is used for the Natural Language processing in this work. We have downloaded a trained Pipe line en_core_web_md which returns a language object that contains all components to process the text. The raw text is the original text that we want to process, and Tokenization is the process of breaking down the raw text into smaller units, called tokens. These can be words, phrases, or even characters. Text Cleaning is the process of removing unwanted characters, symbols, numbers, and other noise from the text. This can include removing HTML tags, punctuation marks, and special characters. POS Tagging that is Part-of-speech (POS) tagging is the process of assigning grammatical tags to each token in the text. This helps to identify the syntactic structure of the text. Stopwords are common words that are often removed from the text during preprocessing, as they do not add much meaning to the text. Examples include "the," "and" "of," etc. The SpaCy library also consists of a stopwords list which by default consists of 326 words, for each word in the input text, if it consists of one of these 326 stopwords, it will be removed. Lemmatization is the process of reducing each word to its base or root

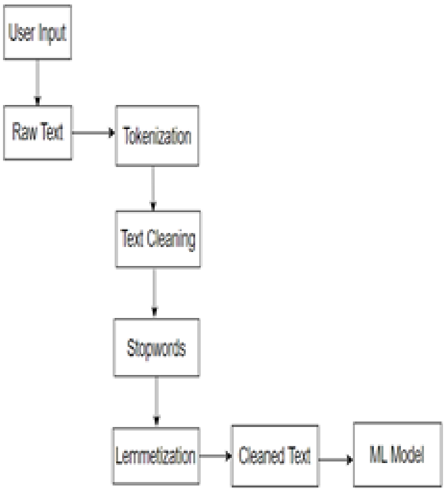


Figure 2. Steps in NLP

form, called a lemma. This helps to reduce the overall vocabulary size and to normalize the text. Cleaned Text is the final output of the NLP pipeline, which is a cleaned and pre-processed version of the original raw text.

Various machine learning algorithms are used in experiments for disease identification based on symptoms. Machine learning algorithms which are used for this experiment and disease prediction process is shown in Figure 3.

Machine learning algorithms which is used for this experiment are listed below:

Support Vector Machine (SVM): It is a powerful supervised learning algorithm and separates data into classes by finding the optimal hyperplane that maximally separates the classes. It is effective for handling high-dimensional data and also handles outliers effectively. In this work “rbf” kernel is used.

Random forest (RB): It is a popular ensemble learning algorithm in machine learning. It combines multiple decision trees to create a robust and accurate model. Each tree in the forest is trained on a random subset of the data and features, reducing overfitting and improving generalization. This is good for handling complex datasets, and providing reliable predictions in various domains.

K-nearest neighbors (KNN): It operates by assigning a new data point to the majority class of its K nearest neighbors in the feature space. It can handle both classification and regression tasks, but it is particularly popular for classification. KNN is easy to implement and understand but the result depends on how parameter K is set (41, 42). Here K=3 used. Algorithm can be computationally expensive for large datasets since it requires calculating distances between data points.

Naïve Bayes (NB): It is a classification algorithm based on Bayes' theorem and the assumption of independence among input features. It is widely used in text classification and spam filtering. It is particularly suitable for large datasets and works well even with limited training data. Used in many real-world applications due to its simplicity and effectiveness. Gaussian Naive Bayes is used in this experiment.

Logistic regression (LR): It is a statistical model used for classification and it predicts the probability of an instance belonging to a specific class Logistic regression is widely used in various domains, including finance, healthcare, and social sciences. In this work class weight balanced parameter is used.

The user inputs their symptoms or medical condition into the system. Here it is in the form of text. In Data preprocessing the input data is pre-processed and cleaned to remove inconsistencies, and irrelevant information. The pre-processed data is the set of symptoms along with their correlated symptoms of threshold more than 0.8 are then fed into machine learning models for classification. Including correlated features along with the given set of symptoms improves prediction accuracy, reduces overfitting, and enhances feature importance. The ML models generate an output that provides a diagnosis or recommendation for the user. The disease which is predicted the maximum that is by the most algorithms is taken into consideration. This may include the likelihood of a particular disease or other relevant information. The system provides the output to the user, and the process is complete.

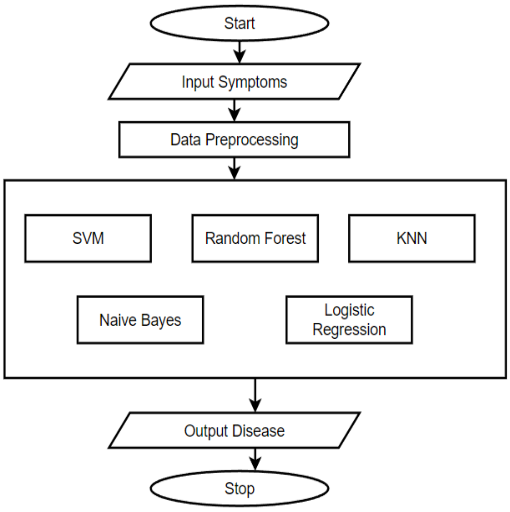


Figure 3. Prediction of diseases using machine learning models

4. RESULTS AND DISCUSSIONS

Natural Language Toolkit (NLTK) is used for implementation, and it has text processing libraries for sentence detection, tokenization, lemmatization, stemming and parsing. Based on the symptoms, most accurate algorithms will be used to classify and using a simple majority the Chatbot must output the diagnosis. At least 4 symptoms must be needed by the Chatbot for the diagnosis to be accurate.

4. 1. Question and Answer Forum

The user initially has to login to the website using email. Login can be done through google authentication or by email id and password entered manually. Users can register using email and password for the first time. User is directed to the main page that is the app component in redux. Different components present are:

Post component: the questions asked by all the users are each wrapped in a post component. Buttons are provided with each post to open a modal to answer the query and another which takes the answer component.

Answer component: displays all the answers or discussions on the particular query.

User component: displays all the questions asked by the user logged in.

Search Component: Displays questions with the keyword provided by the user.

Navbar Component: Contains links to user, search, Chatbot, logout options.

Sidebar Component: Contains various specialization in medical fields on clicking on each will take to the news component which displays latest news on the topic. The news is extracted using the news api.

All the data are stored in the firebase database. Adding, extracting of data is done dynamically from the firestore.

Login Page:

Login page is a component rendered by React which is accompanied with its own stylings. There are two options to login and register which is shown in Figure 4, i.e.,

Sign in with email and password.

Sign in with Google Authentication.

This is an answer forum where users can post questions, and other users or doctors can provide relevant answers. The purpose of this forum is to facilitate the exchange of information and knowledge on a variety of topics, including health, wellness, medicine, and other related subjects. Figure 5 shows Q&A forum.

Once a user clicks on add a question, the user can post a query in the forum which is shown Figure 6. Also, the user can attach any image or link along with the question while submitting it.

Figure 7 shows all the questions posted by the user and the recently posted questions appears on the top.

4. 2. Chatbot Results

The Dr.Bot a disease prediction Chatbot interacts dynamically with the user. The user needs to provide the symptoms they face one after the other and then they get the information about the disease they would probably be facing. The conversation starts by the Chatbot instructing the users the response they need to provide for the prediction to be generated.

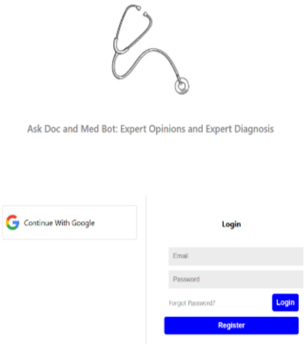
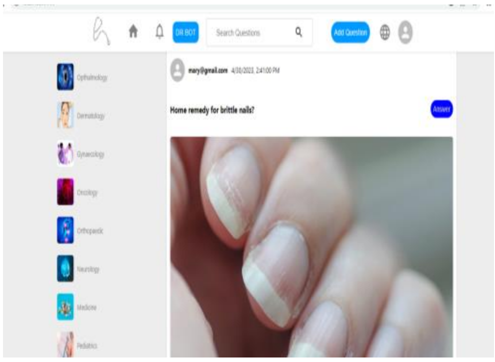


Figure 4. Login and Register page



mary@gmail.com 4/30/2023, 2:41:00 PM

Home remedy for brittle nails?

Figure 5. Question and Answer Forum

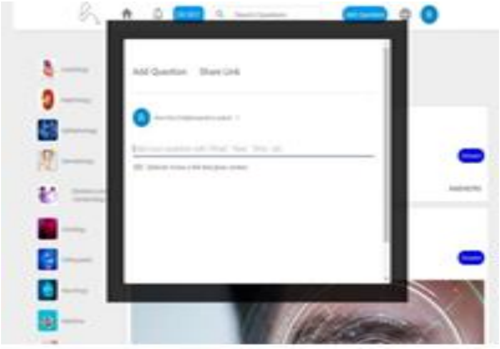
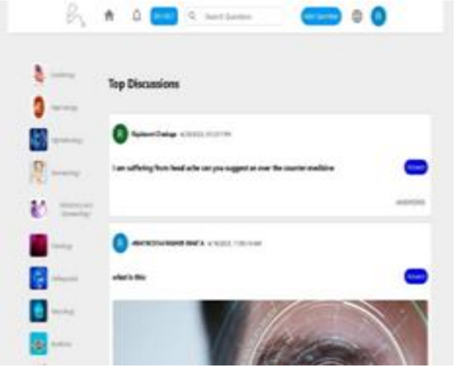


Figure 6. Adding a Question



I am suffering from head ache can you suggest an over the counter medicine

what is this

Figure 7. Posted Questions

The reply from the user cannot be used directly by the Chatbot form machine learning prediction. That is the input must be converted into other form accepted by the ML model. For each input from the user the response is Preprocessed: removal of stop words, pronouns are removed e.g., rash in skin is processed to rash skin.

Finding syntactic similarity: The user might sometimes enter a full sentence with the input symptom in it. For e.g.: I am suffering from muscle pain. The Jaccard set is used to find the intersection between the response and the set of all symptoms in the dataset. If more than one symptom has an intersection with the input provided the list of possible symptoms is provided to the user to select one from the list. If no intersection is found the next step continues.

Jaccard Similarity Between two sets i.e.

X and Y is J (X, Y) = (X∩Y) / (XUY)

If the similarity index is 1 for any symptom in the dataset (i.e., both A and B are the same) that symptom is considered for prediction. If no similarity index of 1 is present with any of the symptoms, all other symptoms which have a similarity score other than 0 are listed and the user is asked to choose the symptom which they are suffering from; e.g., if the user enters pain, all the symptoms in dataset with pain in it is matched and these symptoms have partial similarity with the one user entered (pain in this case)

If all the symptoms have 0 Jaccard similarity with the one user entered, fuzzy string matching is done of the pre-processed user response with all symptoms in dataset. Fuzzy string matching is the technique of finding strings that match with a given string partially and not exactly. A fuzzy similarity score is assigned to the comparison with every symptom, and whichever has the highest similarity score and is above threshold similarity is considered. This is to consider the condition when the user has entered a wrong spelling for a symptom, but the fuzzy string gives a good similarity if not too many spelling errors are there. Example if a user has entered ‘chils’ instead of ‘chills’ a good similarity score is given by the algorithm since they are only one letter apart.

If any of the symptoms do not fuzzy match with the user input, then semantic similarity is found between the input and every symptom.

Finding semantic similarity: If the input response does not match with any symptom in corpus, semantic search is done for the response. The semantic search is done using the all-MiniLM-L6-v2 model of sentence transformer from hugging face. Sentence transformers are models which are pre-trained using a large dataset to find semantic similarity or gives a similarity scores based on their meaning and context. The model assigns similarity score for the two sets provided. In our work we consider a threshold of 0.7 the symptom with similarity passing the threshold is taken into consideration.

If the user enters ‘ache in head’ instead of ‘headache’, the sentence transformer assigns a good similarity score to the above two words since they mean the same thing.

If the response does not meet any of the above criteria, an error message is sent to the user to input again.

A minimum of three symptoms are required for accurate prediction.

The symptoms are then passed to the ML model which evaluates and predicts the most probable disease. Along with the disease, its description and precautions are also provided to users if required. As we know the disease predicted by ML model has the possibility of being a misdiagnosis, the Chatbot after predicting its output always states to please consult a doctor for best opinion. It indicates that the chatbot is just making its prediction for the possible set of symptoms and is always better to reach out to a doctor if the predicted output is something serious. It reiterates the fact that it is just an AI doctor.

All the algorithms used in the project give accuracy above 90% for the test data. For the input symptoms entered by the user the output is accurate when the symptoms are related. In case of unrelated symptoms, the model may not give the best result. All algorithms are applied to test data and out which algorithm gives better results that is considered for final display.

Data Pre-processing Machine Learning Algorithms:

The data is from an open-source site Kaggle[1]. This data has 133 columns and 4920 rows. Each column is a symptom except the last one which represents the prognosis (Class label). Each cell consists of 0 or 1 indicating if the symptom was present or not. If a cell consists of 0 it indicates that the symptom was not seen in that instance. In data preprocessing null values are removed, duplicate rows are removed and few symptoms which won’t give much weightage to overall disease prediction are also removed. We then divide the data set into training data set and testing data set where 75% of the data is used for training the model and 25% is used as the testing data. The Correlation matrix in Figure 8 which consists of 132 rows and 132 columns is hard to understand because of the sheer volume of the symptoms. Selection of highly relevant features is very important to enhance effectiveness of algorithm (43). The code filters out the symptoms which have a correlation more than 0.5 to understand what symptoms are related to each other. For example, the main symptoms of the disease AIDS are if a person was engaging in extramarital contact, had muscle wasting and there occurred patches in the throat.

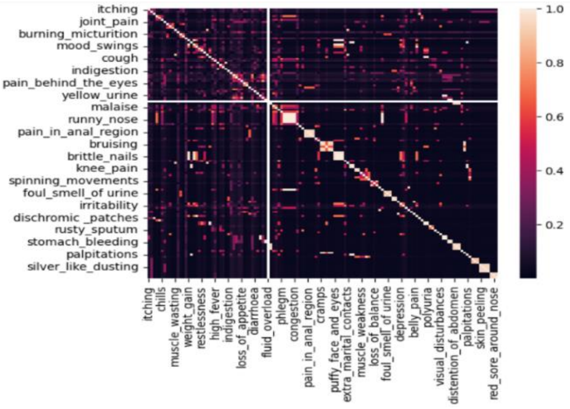


Figure 8. Correlation Matrix

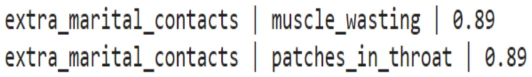
---

[1] https://www.kaggle.com/datasets/itachi9604/disease-symptom-description-dataset?select=dataset.csv

In Figure 9 we see that extra marital contact is related to muscle wasting and patches in throat with the absolute value 0.89. In Figure 10 we see that phlegm is a common symptom with two diseases, with Common Cold and as well as Malaria with different Correlation values. The more the symptoms are correlated to each other, more the probability that the disease predicted is accurate.

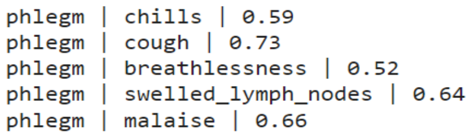
Machine learning algorithms like SVM, RF, KNN, NB and LR algorithm results are evaluated using statistical measures, such as accuracy, precision, recall, and F1-score, to quantify the performance of the chatbot. From the above results observed that the random forest algorithm performs better than all other algorithms for this data (Table 1).

Figure 11 shows results of prediction i.e., common cold using machine learning. To predict diseases, the Chatbot may ask the user a series of questions about their symptoms, medical history, lifestyle habits, and other relevant information. Here we have taken two examples where the Chatbot has predicted diseases that are common cold (Figure 12).



```
extra_marital_contacts | muscle_wasting | 0.89
extra_marital_contacts | patches_in_throat | 0.89
```

Figure 9. Example 1 for Correlation Matrix



```
phlegm | chills | 0.59
phlegm | cough | 0.73
phlegm | breathlessness | 0.52
phlegm | swelled_lymph_nodes | 0.64
phlegm | malaise | 0.66
```

Figure 10. Example 2 for Correlation Matrix

TABLE 1. Evaluation of machine learning algorithm

	SVM	RF	KNN	NB	LR
Accuracy (%)	90.79	97.37	90.79	96.05	94.74
Precision	0.91	0.98	0.89	0.98	0.97
Recall	0.91	0.97	0.91	0.96	0.95
F1-score	0.90	0.97	0.89	0.97	0.94

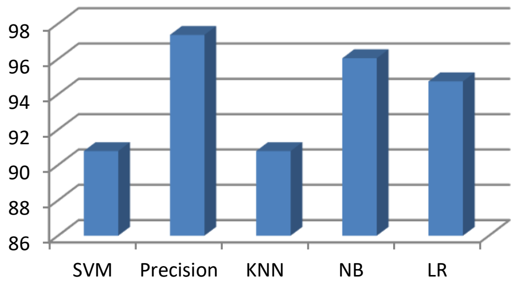
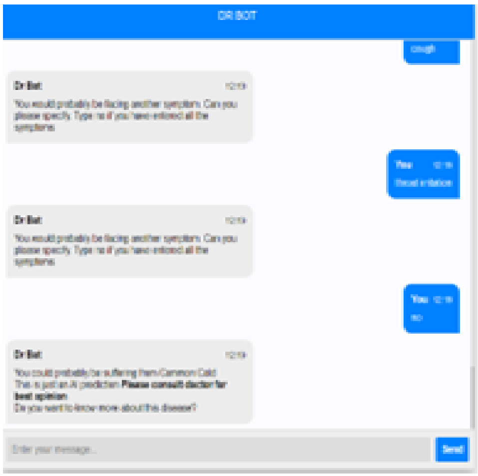


Figure 11. Representation of accuracy of ML model



cough

Dr Bot 12:19
You would probably be facing another symptom. Can you please specify. Type no if you have entered all the symptoms

You 12:19
throat irritation

Dr Bot 12:19
You would probably be facing another symptom. Can you please specify. Type no if you have entered all the symptoms

You 12:19
no

Dr Bot 12:19
You could probably be suffering from Common Cold
This is just an AI prediction. **Please consult doctor for best opinion**
Do you want to know more about this disease?

Enter your message...

Figure 12. Chatbot predicting Common Cold

5. CONCLUSION AND FUTURE SCOPE

In this paper, the disease prediction Chatbot engages with users in real-time to input disease symptoms, and provides a response based on both syntactic and semantic similarity. In syntactic search if input response does not match with any symptom in the corpus, semantic search is done for response. Semantic search seeks to improve search accuracy by understanding the content of the search query and for semantic search all-miniLM -L6v2 sentence transformer model is used. A similarity score threshold of 0.7 is used in this study and if criteria not met means the model asks the user to enter more symptoms. Minimum three symptoms are required to get for better prediction. The symptoms are then passed to the ML model which evaluates and predicts the most probable disease. The Chatbot takes the input as symptoms and predicts the disease for it. Also gives the description of the disease if required. Support vector machine, random forest, K-nearest neighbors, naive bayes and logistic regression algorithms are used for disease prediction, where random forest outperforms with 97.37% accuracy. All algorithms are applied on new unknown test data and out which disease which is predicted by the majority of the algorithms that are taken into consideration and displayed for the user. To enhance the efficiency of the result fuzzy string-matching technique is used. The backend of the Chatbot has been designed where data preprocess of the input from the user is pre-processed to extract only the required symptoms. The frontend development for the login and registration pages, as well as the design of the Chatbot, were completed for the web platform that allows users to post questions and answers. Integration of the Chatbot with the frontend and the backend of the question-and-answer platform are done. In this platform where prediction can be further discussed, a patient can be directed to the actual doctor and get a second opinion there itself.

In future additional features for Question-and-Answer Forum like upvote, downvote etc., can be added to the developed model to make it more interactive. The presence of more upvotes on an answer will make it known to users that the clarity of the answer is good, implying that it is accurate and concise. It will also mean that the answer was useful and helpful to the users. The presence of downvotes would imply that the answers were difficult to understand and have been out of topic to the question asked. The upvotes and downvotes can later be used to filter out the answers there by helping the users access the most relevant answer to their question if it has already been asked before. More diseases and symptoms can be added to the existing dataset to make it more accurate. If addition of data results in more symptoms (features) without new examples (rows), there may be chances of overfitting the training data. But overall, the more assortment of examples we have, our model becomes more general and thus decreases the generalization error. While adding more data there are chances of duplication of the data and inaccurate information creeping into our data set.

7. REFERENCES

1. Michalski RS, Carbonell JG, Mitchell TM. Machine learning: An artificial intelligence approach: Springer Science & Business Media; 2013.
2. Liddy ED. Natural language processing. 2001.
3. Ahangari M, Sebti A. A Hybrid Approach to Sentiment Analysis of Iranian Stock Market User's Opinions. International Journal of Engineering, Transactions C: Aspects,. 2023;36(3):573-84. 10.5829/IJE.2023.36.03C.18
4. Harimi A, Esmaileyan Z. A database for automatic Persian speech emotion recognition: collection, processing and evaluation. International Journal of Engineering, Transactions A: Basics,. 2014;27(1):79-90. 10.5829/idosi.ije.2014.27.01a.11
5. Sadjadi S, Mashayekhi H, Hassanpour H. A two-level semi-supervised clustering technique for news articles. International Journal of Engineering, Transactions C: Aspects,. 2021;34(12):2648-57. 10.5829/IJE.2021.34.12C.10
6. George AS, George AH, Baskar T, Martin AG. Human Insight AI: An Innovative Technology Bridging The Gap Between Humans And Machines For a Safe, Sustainable Future. Partners Universal International Research Journal. 2023;2(1):1-15. https://doi.org/10.5281/zenodo.7723117
7. Lalwani T, Bhalotia S, Pal A, Rathod V, Bisen S. Implementation of a Chatbot System using AI and NLP. International Journal of Innovative Research in Computer Science & Technology (IJIRCST) Volume-6, Issue-3. 2018. http://dx.doi.org/10.2139/ssrn.3531782
8. Athota L, Shukla VK, Pandey N, Rana A, editors. Chatbot for healthcare system using artificial intelligence. 2020 8th International conference on reliability, infocom technologies and optimization (trends and future directions)(ICRITO); 2020: IEEE. 10.1109/ICRITO48877.2020.9197833
9. Turing AM. Mind. Mind. 1950;59(236):433-60.
10. Weizenbaum J. ELIZA—a computer program for the study of natural language communication between man and machine. Communications of the ACM. 1966;9(1):36-45. 10.1145/365153.365168
11. AbuShawar B, Atwell E. ALICE chatbot: Trials and outputs. Computación y Sistemas. 2015;19(4):625-32. 10.13053/cys-19-4-2326
12. Xu L, Sanders L, Li K, Chow JC. Chatbot for health care and oncology applications using artificial intelligence and machine learning: systematic review. JMIR cancer. 2021;7(4):e27850. 10.2196/27850
13. Saengrith W, Viriyavejakul C, Pimdee P. Problem-Based Blended Training via Chatbot to Enhance the Problem-Solving Skill in the Workplace. Emerging Science Journal. 2022;6:1-12. 10.28991/ESJ-2022-SIED-01
14. Mendonça YV, Naranjo PGV, Pinto DC. The Role of Technology in the Learning Process. Emerging Science Journal. 2022;6(Special Issue):280-95. 10.28991/ESJ-2022-SIED-020
15. Roshani S, Coccia M, Mosleh M. Sensor technology for opening new pathways in diagnosis and therapeutics of breast, lung, colorectal and prostate cancer. medRxiv. 2022:2022.02. 18.22271186. 10.28991/HIJ-2022-03-03-010
16. Hamidi H, Daraee A. Analysis of pre-processing and post-processing methods and using data mining to diagnose heart diseases. International Journal of Engineering, Transactions A: Basics 2016;29(7):921-30. 10.5829/idosi.ije.2016.29.07a.06
17. Ye EZ, Ye EH, Bouthillier M, Ye RZ. DeepImageTranslator V2: analysis of multimodal medical images using semantic segmentation maps generated through deep learning. bioRxiv. 2021:2021.10. 12.464160. 10.28991/HIJ-2022-03-03-07
18. Manjurul Ahsan M, Siddique Z. Machine learning based disease diagnosis: A comprehensive review. arXiv e-prints. 2021:arXiv: 2112.15538.
19. Shaji SP, editor Predictionand diagnosis of heart disease patients using data mining technique. 2019 international conference on communication and signal processing (ICCSP); 2019: IEEE. 10.1109/ICCSP.2019.8697977
20. Keniya R, Khakharia A, Shah V, Gada V, Manjalkar R, Thaker T, et al. Disease prediction from various symptoms using machine learning. Available at SSRN 3661426. 2020. http://dx.doi.org/10.2139/ssrn.3661426
21. Ferjani MF. Disease Prediction Using Machine Learning. Bournemouth, England: Bournemouth University. 2020. 10.13140/RG.2.2.18279.47521
22. Magoulas GD, Prentza A. Machine learning in medical applications. Advanced course on artificial intelligence: Springer; 1999. p. 300-7.
23. Jain R, Chotani A, Anuradha G. Disease diagnosis using machine learning: A comparative study. Data Analytics in Biomedical Engineering and Healthcare: Elsevier; 2021. p. 145-61.
24. Pingale K, Surwase S, Kulkarni V, Sarage S, Karve A. Disease prediction using machine learning. International Research Journal of Engineering and Technology (IRJET). 2019;6(12):831-3.
25. Bharti U, Bajaj D, Batra H, Lalit S, Lalit S, Gangwani A, editors. Medbot: Conversational artificial intelligence powered chatbot for delivering tele-health after covid-19. 2020 5th international conference on communication and electronics systems (ICCES); 2020: IEEE. 10.1109/ICCES48766.2020.9137944
26. Tjiptomongsoguno ARW, Chen A, Sanyoto HM, Irwansyah E, Kanigoro B. Medical chatbot techniques: a review. Software Engineering Perspectives in Intelligent Systems: Proceedings of 4th Computational Methods in Systems and Software 2020, Vol 1 4. 2020:346-56. 10.1007/978-3-030-63322-6_28
27. Reshma R, editor An Improved Chatbot for Medical Assistance using Machine Learning. 2022 International Conference on Inventive Computation Technologies (ICICT); 2022: IEEE. 10.1109/ICICT54344.2022.9850470

28. Prayitno PI, Leksono RPP, Chai F, Aldy R, Budiharto W, editors. Health Chatbot Using Natural Language Processing for Disease Prediction and Treatment. 2021 1st International Conference on Computer Science and Artificial Intelligence (ICCSAI); 2021: IEEE. 10.1109/ICCSAI53272.2021.9609784
29. Ayanouz S, Abdelhakim BA, Benhmed M, editors. A smart chatbot architecture based NLP and machine learning for health care assistance. Proceedings of the 3rd international conference on networking, information systems & security; 2020.
30. Caldarini G, Jaf S, McGarry K. A literature survey of recent advances in chatbots. Information. 2022;13(1):41. https://doi.org/10.3390/info13010041
31. Kumar R, Ali MM. A review on chatbot design and implementation techniques. Int J Eng Technol. 2020;7(11).
32. Chaudhary J, Joshi V, Khare A, Gawali R, Manna A. A Comparative Study of Medical Chatbots. International Research Journal of Engineering and Technology (IRJET). 2021;8(02).
33. Vasileiou MV, Maglogiannis IG. The Health ChatBots in Telemedicine: Intelligent Dialog System for Remote Support. Journal of Healthcare Engineering. 2022;2022. https://doi.org/10.1155/2022/4876512
34. Shaikh A, More D, Puttoo R, Shrivastav S, Shinde S. A survey paper on chatbots. International Research Journal of Engineering and Technology (IRJET). 2019;6(04):2395-0072.
35. PhaniRaghavaa B, Kumarb SA. An Improved Chatbot for Predicting Disease and Medicines Using Natural Language Processing with Fuzzy Logic. Advances in Parallel Computing Algorithms, Tools and Paradigms. 2022;41:258. 10.3233/APC220035
36. Amer E, Hazem A, Farouk O, Louca A, Mohamed Y, Ashraf M, editors. A proposed chatbot framework for COVID-19. 2021 International Mobile, Intelligent, and Ubiquitous Computing Conference (MIUCC); 2021: IEEE. 10.1109/MIUCC52538.2021.9447652
37. Kumar A, Sharma GK, Prakash U. Disease prediction and doctor recommendation system using machine learning approaches. International Journal for Research in Applied Science & Engineering Technology (IJRASET). 2021;9:34-44. https://doi.org/10.2214/ijraset.20221.36234
38. Tamizharasi B, Livingston LJ, Rajkumar S, editors. Building a medical chatbot using support vector machine learning algorithm. Journal of Physics: Conference Series; 2020: IOP Publishing. 10.1088/1742-6596/1716/1/012059
39. Giansanti D. The Chatbots Are Invading Us: A Map Point on the Evolution, Applications, Opportunities, and Emerging Problems in the Health Domain. Life. 2023;13(5):1130. https://doi.org/10.3390/life13051130
40. Wilson L, Marasoiu M. The development and use of chatbots in public health: scoping review. JMIR human factors. 2022;9(4):e35882. 10.2196/35882
41. Fooladi S, Farsi H, Mohamadzadeh S. Segmenting the lesion area of brain tumor using convolutional neural networks and fuzzy k-means clustering. International Journal of Engineering, Transaction B: Applications. 2023;36(8):1556-68. 10.5829/IJE.2023.36.08B.15
42. Pourbahrami S, Balafar MA, Khanli LM, Kakarash ZA. A survey of neighborhood construction algorithms for clustering and classifying data points. Computer Science Review. 2020;38:100315. https://doi.org/10.1016/j.cosrev.2020.100315
43. Kakarash ZA, Mardukhia F, Moradi P. Multi-label feature selection using density-based graph clustering and ant colony optimization. Journal of Computational Design and Engineering. 2023;10(1):122-38. https://doi.org/10.1093/jcde/qwac120

Persian Abstract

چکیده

دسترسی به موقع به مراقبت های بهداشتی برای حفظ استاندارد بالای زندگی بسیار مهم است. با این حال، دریافت مشاوره پزشکی می تواند دشوار باشد، به ویژه برای کسانی که در مناطق دور افتاده زندگی می کنند یا در طول یک بیماری همه گیر که مشاوره حضوری همیشه امکان پذیر نیست. توانایی تشخیص دقیق بیماری ها برای درمان موثر ضروری است و پیشرفت های تکنولوژیکی اخیر یک راه حل بالقوه ارائه می دهد. یادگیری ماشینی (ML)و پردازش زبان طبیعی (NLP)برنامه‌های رایانه‌ای را قادر می‌سازد تا زبان انسان را بفهمند و ویژگی‌های مورد نظر را از پاسخ‌ها استخراج کنند و امکان تعامل انسان‌مانند با کاربران را فراهم کنند. با استفاده از این فناوری‌ها، متخصصان مراقبت‌های بهداشتی به طور بالقوه می‌توانند مشاوره‌های پزشکی در دسترس و کارآمدتری را برای افراد، صرف‌نظر از موقعیت مکانی آن‌ها، ارائه دهند. مفهوم ایجاد یک پلت فرم آنلاین است که در آن کاربران می توانند سوالات مربوط به پزشکی را بپرسند و پاسخ هایی را هم از متخصصان پزشکی و هم از کاربران دیگر دریافت کنند. این پلتفرم دارای یک Chatbot پزشکی است که از تکنیک های پیشرفته ML برای تجزیه و تحلیل علائم ارائه شده توسط کاربر و ارائه تشخیص اولیه بیماری و اطلاعات مرتبط قبل از مشورت با پزشک استفاده می کند. این چت بات پیش بینی بیماری به صورت پویا با کاربران وارد تعامل می شود تا علائم بیماری ها را وارد کند و بر اساس شباهت نحوی و معنایی پاسخ داده می شود. در این کار، امتیاز آستانه تشابه ۰.۷ حفظ شده است. K-نزدیکترین همسایه، جنگل تصادفی، ماشین بردار پشتیبان، سیسنم ساده و الگوریتم های رگرسیون لجستیک برای پیش بینی بیماری بر اساس علائمی که کاربران با آن مواجه هستند استفاده می شود. شباهت نحوی، تطبیق رشته فازی و شباهت معنایی با استفاده از مدل all-MiniLM-L6-v2 برای بهبود کارایی نتیجه استفاده می شود.

IJE TRANSACTIONS B: Applications Vol. 37 No. 02, (February 2024) 323-340

International Journal of Engineering

Journal Homepage: www.ije.ir

Comprehensive Framework of Influential Factors on Innovation Ecosystem Resilience: Using Meta-Synthesis and Structural Equation Modelling

S. Abdi[a], M. Yazdani*[b], E. Najafi[a]

[a] *Department of Industrial Engineering, Science and Research branch, Islamic Azad University, Tehran, Iran*
[b] *Department of Industrial Engineering, Qazvin Branch, Islamic Azad University, Qazvin, Iran*

PAPER INFO

Paper history:
Received 20 August 2023
Received in revised form 24 September 2023
Accepted 01 October 2023

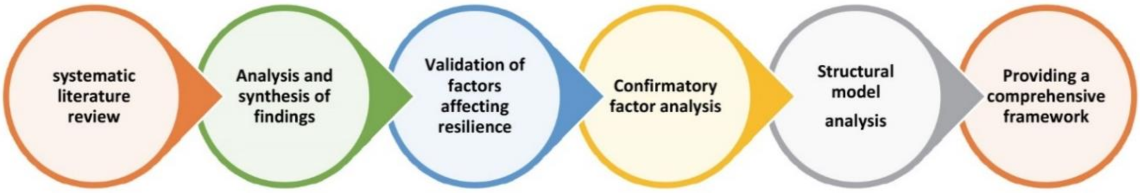
Keywords:
Innovation Ecosystem
Resilience
Meta Synthesis
Confirmatory Factor Analysis
Structural Equation Modelling

ABSTRACT

Resilience of the innovation ecosystem as a driving force of knowledge-based economies, provides relative stability against environmental disruptions. Currently, finding a comprehensive framework of factors influencing innovation ecosystem resilience is a major concern of policymakers to effectively select policies of resilience improvement. This research analyzes and presents a comprehensive framework of factors influencing innovation ecosystem resilience by using meta-synthesis approach, confirmatory factor analysis and structural equation modelling. These factors include Adaptability, Innovation Management, Recovery Capability, Culture, Resource, Robustness, Strategic Planning, and Vulnerability. In this paper, Iranian Power Innovation Ecosystem is considered as a case study. The computational results indicate that vulnerability and adaptability are the most influential factors on innovation ecosystem resilience, while recovery capacity and resiliency culture are less impactful factors. The innovative aspects of this study include the use of meta-Synthesis method for systematic review, content analysis, and categorization of influential factors, as well as the presentation of a comprehensive framework based on factor analysis and structural modelling. The findings of this research assist innovation ecosystem policymakers in evaluating various factors and planning accordingly to achieve their desired goals based on the aforementioned framework.

***doi**: 10.5829/ije.2024.37.02b.09*

Graphical Abstract



1. INTRODUCTION

Today, it is increasingly important for organizations to adjust their innovation strategy from organization-based to ecosystem-based (1). Due to dynamics social complexity of innovation and concept of complex adaptive systems; policymakers change their mind from an innovation system to an innovation ecosystem (2). Innovation ecosystem is a set of innovation entities which interact in a dynamic environment (3) that enables collective work to empower knowledge flow, supporting technology development, and creating innovation (4). Due to significant role of innovation as a source of competitive advantage, policymakers are seeking solutions to improve the performance of the innovation ecosystem.

The resilience of an innovation ecosystem refers to its capacity to absorb disturbances, and reorganize throughout undergoing changes, to maintain the core functions. This resilience significantly influences the ecosystem's overall performance (5). Unsuitable resilience assessment can lead to weak strategies to

---

*Corresponding Author's Email: Mehdi_yazdani2007@yahoo.com (M. Yazdani)

Please cite this article as: Abdi S, Yazdani M, Najafi E. Comprehensive Framework of Influential Factors on Innovation Ecosystem Resilience: Using Meta-Synthesis and Structural Equation Modelling, International Journal of Engineering, Transactions B: Applications. 2024;37(02):323-40.

reduce the severity effects of possible disruption and inappropriate adaptation to environmental conditions (5) Resilient systems are characterized by system characteristics that affect different components of resilience, such as robustness, redundancy, reliability, readiness, vulnerability, sustainability, and adaptive capacity (5). Bai and Li (6) emphasized that the stability and development of the innovation ecosystem, are enhanced by improving internal structure, better adapting to external environment changes, and having the ability to restore after damage.

As stated, the main reason for this article's focus on research about the resilience of the innovation ecosystem is to help preserve and improve its performance in different situations by making it resilient against various crises. This makes the knowledge-based economy in the countries to enjoy stability and prosperity and improve the level of well-being of human societies.

Resilience evaluation methods vary depending on evaluation objectives and the characteristics of the study system. Hosseini et al. (7) categorized them into two quantitative and qualitative methods. The concept of resilience requires a comprehensive framework of factors that influence on resilience, as well as an integrated assessment for a better understanding, which we have not noticed in previous provided researches (8). So, we seek to develop a comprehensive framework of factors affecting the resilience of the innovation ecosystem by using meta-synthesis and structural equation modeling technique in our study. Hosseini et al. (7) identified economical, engineering, social and organizational as four main area of resilience. Meanwhile other researchers such as Nylund et al. (9) define innovation ecosystems as a loose network of organizations and individuals collaborate and evolve together to create value through innovation. Also, according to Xiao and Cao (10) the resilience of a whole system is achieved through its components resiliency. Therefore, it is essential for all actors of the ecosystem including different organizations, to be resilient to achieve the overall resilience of the ecosystem. with this in mind, our literature review focus on the system resilience and organization resilience. Then we analyzed and presented a comprehensive framework of factors influencing innovation ecosystem resilience by using meta-synthesis approach, confirmatory factor analysis (CFA) and structural equation modelling (SEM). Our statistical population consisted of actors of the *Iranian Power Innovation Ecosystem.*

In general, the contribution of this paper includes the presentation of a comprehensive framework of influential factors on innovation ecosystem resilience by using meta-synthesis and structural equation modelling.

The rest of the paper is structured as follows: Section 2 presents a research background, it covers past studies in the field of innovation ecosystem resilience, different systems and organizations, and related factors. Section 3 introduces problem definition and research method. Section 4 contains research findings include primary framework, CFA, SEM analysis and comprehensive framework of the factors affecting the resilience of the innovation ecosystem. Finally, conclusion and some suggestions for further research are presented in section 5.

2. RESEARCH BACKGROUND

Walker et al. (11) stated that resilience has four components: elasticity (the maximum amount a system can change before losing its ability for recovery), resistance, instability and panarchy. Herrera and Kopainsky (12) combined system dynamics (SD) and participatory approaches to assess resilience.

Egli et al. (13) investigated resilience and agent-based models (ABMs) in ecological and social-ecological systems by reviewing the literature. Bruneau et al. (14) presented a conceptual framework for defining the resilience of communities after an earthquake and quantitative measures of resilience. Sauser et al. (15) investigated the resilience of SMEs after a disaster event. In their study, resilience is defined as an adaptive capacity of SMEs to meet and achieve priorities and objectives in order to absorb and limit disruptions while maintaining service continuity. Henry and Ramirez-Marquez (16) proposed general criteria and a quantitative method to evaluate the resilience of systems. Sweetapple et al. (17) presented a framework to plan a reliable and resilient system through the integration of multi-objective optimization and reliability, robustness, and resilience assessment. Ahmadi et al. (18) proposed characteristics of prediction, absorption, adaptation and recovery for energy system resilience modelling. Tran et al. (19) provided a flexible conceptual framework that can be applied to a variety of systems. Biddle et al. (20) believed that most of the literature on the previous topic only deals with certain aspects such as absorptive and adaptive capacities in relation to the resilience of systems; while the rightfulness of institutions and transformative capacity have rarely been considered. Zhu et al. (21) studied the assessment of the infrastructure systems resilience using eight factors including: vulnerability, predictability, redundancy, adaptive capacity, speed, resourcefulness, interactions, and learning culture. Yu et al. (22) introduced Community Resilience Cost Index (CRCI) as a method for resilience assessment.

Burnard and Bhamra (23) reviewed the challenges of organizational resilience and declare that in order for an organization to be considered resilient, both active resilience (active participation and adjustment of a system in relation to change) and passive resilience (the ability to resist or absorbing disorders) are required.

Hillmann and Guenther (24) identified concepts such as awareness, sustainability, change, growth and performance for organizational resilience. Ma et al. (25) proposed an organizational resilience integrated model including three dimensions of cognitive, behavioural and contextual resilience. On the other hand, Hamel and Valikangas (26) determined organizational resilience challenges into four categories: cognitive, strategic, political and ideological. Chen et al. (27) used grounded theory to extract the main characteristics of organizational resilience. They indicated that organizational resilience includes five dimensions: financial, cultural, strategic, communication and learning resilience. But in other research, Annarelli et al. (28) considered adaptability, reliability, agility, effectiveness, flexibility, recovery level and recovery time. Lee et al. (29) provided a tool includes two main factors of adaptability and planning capacity and consists of thirteen indicators to measure and compare the resilience of organizations. McManus et al. (30) proposed a model called the "Relative Overall Resilience" (ROR), in which relative overall resilience consists of three factors: awareness of the situation, management of key vulnerabilities, and adaptive capacity. Rahi (31) also states the two dimensions of awareness and adaptability that are related to organizational resilience. Xiao and Cao (10) presented a theoretical model of organizational resilience. In their model an organization must become resilient at the individual, group and organizational levels; and the factors for each one is different. Sanchis and Poler (32) presented a quantitative approach to increase the resilience of the organization with a dynamic planning approach. Aleksić et al. (33) evaluated the potential of organizational resilience in SMEs in process industries using fuzzy mathematical modelling. Tong et al. (34) identified five key factors empower the resilience of the innovation ecosystem of high-tech companies, including resilience thinking, environmental uncertainties, tolerance threshold, evolutionary capacity and hidden resources. Roundy et al. (35) have shown that the interaction between ecosystem diversity and cohesion would result in ecosystem resilience.

Based on the literature review, the subject of measuring the resilience of the innovation ecosystem has received less attention, additionally a comprehensive framework of factors affecting the resilience of the innovation ecosystem based on a comprehensive qualitative research method has not been presented yet. Thus, the main objective of our research is to present a comprehensive framework of factors influencing innovation ecosystem resilience by using a systematic literature review, meta-synthesis approach and structural equation modelling. on the other hand, Innovation ecosystems not only vary in their architecture and internal collaboration models (36), but they also differ significantly in terms of actors, governance, investments, and business models. Therefore, the tailoring and adaptation of the resilience factors framework for the innovation ecosystem as a case study for Iran's Power industry highlights the novelty of this research.

3. PROBLEM DEFINITION AND RESEARCH METHOD

This research, through a systematic literature review and a holistic view, seeks to resolve the research gap stated in the background section of the research and answer the following question:

- What factors affect the resilience of the innovation ecosystem and how are they connected to resilience?

To answer the above question and present the results, it is necessary to follow steps shown in Figure 1.

The first and second steps are related to the systematic literature review, information extraction and analysis, combining the findings and extracting the primary factors affecting the resilience of the innovation ecosystem. In the third step, the validation of the extracted factors is done to present the primary framework. The concepts are categorized according to a scientific method and the overall outline of the primary framework is formed. For this purpose, the meta-synthesis method, which provides a systematic approach for researchers by combining past research, discovers new theme; is used. It should be

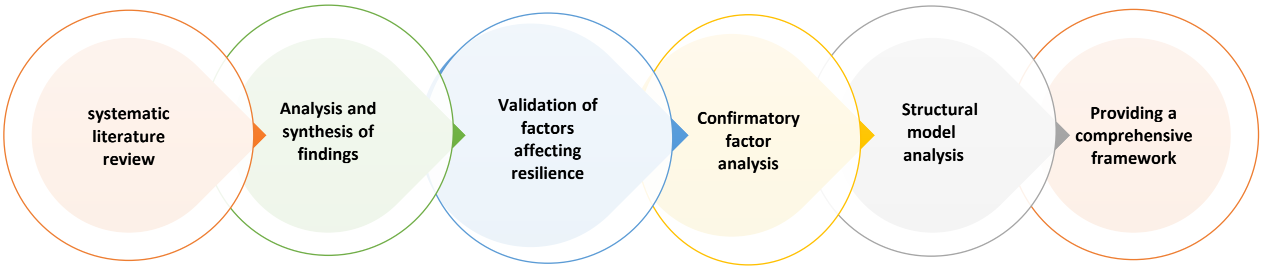


Figure 1. The steps of providing a comprehensive framework

noted that so far, based on our literature review, meta-synthesis approach has not been used in innovation ecosystem resilience studies. Sandelowski and Barroso (37) have presented a seven-step method in connection with the use of meta-synthesis method, which is more general. These steps consist of formulating the review questions; Systematic review of the literature; Selecting appropriate research articles; Extracting the results; Analyzing and synthesizing qualitative findings; validating and presenting findings (primary framework) that will be used in this research as well. It should be noted that after providing the primary framework, using confirmatory factor analysis and then structural model analysis, a comprehensive framework of factors affecting the resilience of the innovation ecosystem is finally presented.

4. FUNDING AND RESULTS

In this section, the findings of research are presented separately in related subsections. First, the primary factors affecting the innovation ecosystem resilience, which is the result of articles content analysis, and also how to validate these findings are presented. Then, in the continuation, the method of using confirmatory factor analysis and structural equation analysis to finalize the observed and latent variables as well as the test of various Hypothesized path is presented. Finally, based on the results, the comprehensive framework of factors affecting the resilience of the innovation ecosystem and additional explanations are presented.

4. 1. Systematic Literature Review In this research, we have taken a comprehensive approach to gather relevant information. The study includes articles from peer reviewed journals and prestigious conferences between 2002 and the first half of 2022 in Scopus and Web of Science. Also, to ensure an inclusive search, we have utilized Google Scholar. According to literature review, the keywords of "*system resilience*", "*ecosystem resilience*", "*organizational resilience*", and "*resilience assessment*" were used to search articles more comprehensively. Vast number of articles (217) were found in the initial research based on keywords. After reviewing the abstract and content of the articles and removing the articles with incomplete information or unrelated to the specified topic, finally we select 33 articles for full review and content analysis.

4. 2. Analysis and Synthesis of Findings Factors affecting the resilience of the innovation ecosystem recognized based on the content analysis of reviewed articles. Each concept identified in the articles was considered as a code, and then codes with similar meanings were grouped into categories. Ultimately, 60 concepts were extracted, leading to the identification of 8 categories including: culture, strategic planning, resources, innovation management, vulnerability, robustness, adaptability and recovery capability as shown in Table 1. Among these factors, adaptability was the most frequently mentioned, while innovation management was the least frequently referred to as shown in Figure 2.

TABLE 1. Extracted concepts and their categorizations

No.	Reference	Concept	Category
1	(13, 26, 29, 38-40)	Culture	Culture
2	(25, 41)	Risk Management Culture	
3	(26)	Relationship Resilience	
4	(13, 26, 41)	Interaction	
5	(38, 39)	Resilience Thinking	
6	(5, 7, 8, 24, 30, 42, 43)	Awareness- Sense Making	
7	(8, 9, 13, 20, 26, 42, 43)	Learning Culture	
8	(44 ,27 ,26)	Strategic Resilience	Strategic Planning
9	(2, 12, 16, 42, 45, 46)	Planning	
10	(38)	Flexibility in Organizational Strategy	
11	(9, 13, 45, 47, 48)	Resourcefulness	
12	(13, 16, 42, 49, 50)	Anticipation	
13	(16, 51)	Efficiency/Effectiveness	

No.	Reference	Concept	Category
14	(52 ,51 ,33 ,29 ,28 ,25)	Information Systems	Resources
15	[(34)	Financial Resources, Intangible Resources	
16	(53 ,46 ,21 ,14 ,7)	Cash Flow	
17	(54 ,28 ,25 ,21 ,18)	Social Capital	
18	(49 ,28)	Talent Diversity ‹Personal Skills	
19	(55 ,6)	Enterprise Diversity	Innovation Management
20	(56 ,55 ,45 ,41 ,34 ,33 ,27 ,15 ,6)	Internal Structural Complexity	
21	(6)	Robust Organizational Structure	
22	(56 ,53 ,34 ,10)	Entrepreneurial Talents	
23	(34 ,6)	Diversity of Innovation Subjects	
24	(6)	Innovation Capital Investment	
25	(6)	Innovation Output	
26	(34)	Technology Stream	
27	(56 ,34)	Redundancy	Vulnerability
28	(55 ,48 ,6)	Vulnerability	
29	(6)	Vulnerability Management	
30	[(6)	Business Environments	
31	(6)	Governance Condition	
32	(58 ,57 ,53 ,49 ,46 ,45 ,43 ,28 ,21 ,14 ,7)	Environmental Uncertainty	Robustness
33	(60 ,59 ,51 ,42 ,28 ,21)	Sustainability	
34	(61 ,30)	Resistance	
35	(56 ,39)	Overall Resilience Strategy	
36	(51)	Robustness	
37	(34)	Tolerance	
38	(62 ,28 ,24 ,11 ,6)	Absorption Capacity	Adaptability
39	(63 ,51 ,23 ,11)	Business Continuity	
40	(16)	Adaptive Capacity, Adaptability	
41	(53 ,46 ,45 ,38 ,17 ,14)	Evolutionary Ability	
42	(34 ,15)	Transformative Capacity	
43	(51 ,20-18 ,7)	Self-Organization	
44	(51)	Readiness	
45	(64 ,61 ,58 ,54 ,51 ,48 ,31-28 ,24 ,21-18 ,15 ,10 ,7)	The Appropriate Response of Organization	
46	(34)	Coping Capacity	
47	(64 ,20)	Improvisation	
48	(58 ,57)	Latitude	
49	(66 ,65 ,60 ,59 ,57 ,51 ,42 ,28 ,23)	Flexibility	
50	(23)	Self-Adjustment	
51	(64 ,24)	Self-Adaptation	
52	(10)	Recovery	Recovery Capability
53	(11)	Recovery Time and Speed	
54	(57 ,49 ,45 ,34 ,28)	Agility	
55	(6)	Recovery Costs	

No.	Reference	Concept	Category
56	(6)	Recovery Level	
57	(67 ,55 ,51 ,50 ,19 ,18 ,16 ,7)	Growth	
58	(70-68 ,60 ,59 ,51 ,49 ,46 ,42 ,28 ,21 ,14)	Performance Level	
59	(49 ,28)	Ability to Reconfigure	
60	(22 ,18)	Availability	

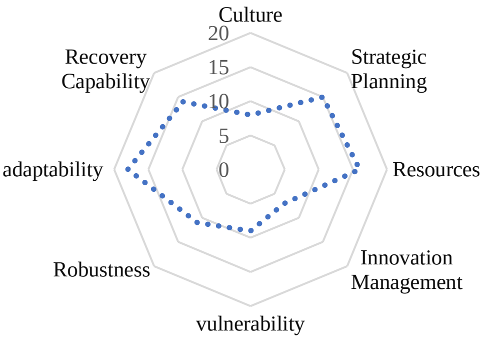


Figure 2. Frequency of extracted factors from articles

4. 3. Validation of Factors Affecting the Resilience of the Innovation Ecosystem

In this step of the research, for validation concepts and their related categories, we used experts' opinion. in a purposeful way nine experts selected from university and industry who were familiar with the power innovation ecosystem. Academic experts must have participated in at least one of the innovative technological projects in power innovation ecosystem, or have published research papers or books in the field of innovation. In connection with industrial experts, having a managerial position in one of the players of the innovation ecosystem, as well as a history of performing activities related to the performance indicators of power innovation ecosystem, have been considered. To assess the importance of each concept, we employed Lawshe's content analysis method (71). Each expert was asked to evaluate the importance of each concept using a triple spectrum consisting of three categories: "*essential*", "*useful but not necessary*" and "*not necessary*".

The Content Validity Ratio (CVR) and the Numerical Average of Judgments (NAJ) for each concept were computed That $CVR = (ne - \frac{N}{2})/\frac{N}{2}$, where " ne " represents the number of experts indicating "essential" and "N" represents the total number of experts. Additionally, we computed the Content Validity Index (CVI) for each category. That $CVI = \sum CVR/Retained\ Numbers,\ and\ "Retained\ Numbers"$ present the number of approved concepts. the attribute which has values greater than 0.78 have been accepted and shown in Table 2.

TABLE 2. CVR, CVI for concepts and their related categories

No.	Concept	Category	CVR	NAJ	CVI
1	Culture	Culture	1.00	2.00	94%
2	Risk Management Culture		0.78	1.89	
3	Relationship Resilience		1.00	2.00	
4	Interaction		1.00	2.00	
5	Resilience Thinking		1.00	2.00	
6	Awareness- Sense Making		1.00	2.00	
7	Learning Culture		0.78	1.89	
8	Strategic Resilience	Strategic Planning	0.78	1.89	87%
9	Planning		0.78	1.89	
10	Flexibility in Organizational Strategy		1.00	2.00	
11	Resourcefulness		0.78	1.89	
12	Anticipation		1.00	2.00	
13	Efficiency		0.78	1.89	
14	Effectiveness		1.00	2.00	

No.	Concept	Category	CVR	NAJ	CVI
15	Information Systems	Resources	1.00	2.00	93%
16	Financial Resources		1.00	2.00	
17	Intangible Resources		0.78	1.89	
18	Cash Flow		0.78	1.89	
19	Social Capital		1.00	2.00	
20	Personal Skills		0.78	1.89	
21	Enterprise Diversity		1.00	2.00	
22	Internal Structural Complexity		1.00	2.00	
23	Robust Organizational Structure		1.00	2.00	
24	Entrepreneurial Talents	Innovation Management	1.00	2.00	91%
25	Diversity of Innovation Subjects		0.78	1.89	
26	Innovation Capital Investment		0.78	1.89	
27	Innovation Output		1.00	2.00	
28	Technology Stream		1.00	2.00	
29	Redundancy	Vulnerability	1.00	2.00	89%
30	Vulnerability		1.00	2.00	
31	Vulnerability Management		1.00	2.00	
32	Business Environments		0.78	1.89	
33	Governance Condition		0.78	1.89	
34	Environmental Uncertainty		0.78	1.89	
35	Sustainability	Robustness	1.00	2.00	94%
36	Resistance		1.00	2.00	
37	Overall Resilience Strategy		1.00	2.00	
38	Robustness		1.00	2.00	
39	Tolerance		0.78	1.89	
40	Absorption Capacity		0.78	1.89	
41	Business Continuity		1.00	2.00	
42	Adaptive Capacity	Adaptability	1.00	2.00	87%
43	Evolutionary Ability		1.00	2.00	
44	Transformative Capacity		0.78	1.89	
45	Self-Organization		1.00	2.00	
46	Readiness		1.00	2.00	
47	The Appropriate Response of Organization		1.00	2.00	
48	Coping Capacity		0.78	2.00	
49	Improvisation		0.78	2.00	
50	Latitude		0.78	2.00	
51	Flexibility		0.78	2.00	
52	Self-Adjustment		0.78	2.00	
53	Self-Adaptation		0.78	2.00	
54	Recovery	Recovery Capability	1.00	2.00	93%
55	Recovery Speed		0.78	1.89	
56	Agility		1.00	2.00	
57	Recovery Costs		0.78	1.89	
58	Recovery Level		1.00	2.00	
59	Growth		0.78	1.89	
60	Performance Level		1.00	2.00	
61	Ability to Reconfigure		1.00	2.00	
62	Availability		1.00	2.00	

According to the results obtained from the meta-synthesis and validation method, eight factors affecting the resilience of the innovation ecosystem include culture, strategic planning, resources, innovation management, vulnerability, robustness, adaptability and recovery capability are discovered that formed the base of primary framework.

4. 4. Confirmatory Factor Analysis As presented in previous section, the primary framework includes eight factors. In this step, based on confirmatory factor analysis, all factors were considered as latent variables. Then, in order to measure and analyze them, a number of observed variables that will form the questionnaire were identified. These observed variables were formulated based on the analysis and Synthesis of findings from the literature review (Section 4.2). Therefore, for each latent variable, a number of observed variables have been recognised; and questionnaire with 39 observed variable (questions) was prepared. Therefore, this researcher-made questionnaire was used to survey a statistical population consisting of all actors of Power Innovation Ecosystem that includes the Deputy of Research of Ministry of Energy, Power Generation, Transmission & Distribution Company, Universities, Power Research Institute, Venture Capitalists, Incubators, and Knowledge-based companies. In this questionnaire 5-point Likert scale (Strongly Disagree=1, Disagree=2, Neutral=3, Agree=4, Strongly Agree=5) used for each question. It is important to note that the questionnaires were collected in both paper and electronic forms.

The process of collecting and analysing data by research made questionnaire was done in two stages:

- First stage: selecting 30 experts familiar with the subject, conducting factor analysis and finalizing the questionnaire
- The second stage: sending the final questionnaire to a larger number of experts, collecting the completed ones and analyze through the structural modelling equations.

In the first stage, 30 experts familiar with the Iranian power innovation ecosystem were selected using the snowball method in a purposeful manner, Subsequently, we distributed the initial questionnaire consisting of 39 observed variables related to 9 latent variables, to these experts. All the questionnaires were collected and utilized in confirmatory factor analysis. In factor analysis, it is determined, which observed variables are suitable to descript a latent variable and which are not. after removing unsuitable observed variables, the model is run again until finally all the remaining observed variables are recognized as appropriate. This research used Partial Least Squares (PLS) method that is a non-parametric multivariate statistical technique that allows comparison between multiple observed variables and multiple latent variables. PLS is one of a covariance-based statistical methods which are often referred to as structural equation modeling that designed to cope with problems in data specifically, small datasets, missing values and multicollinearity.

The data of the collected questionnaires were entered into the *SmartPLS* software. The standard deviation associated with all observed variables was examined if it is equal to zero, that variable was removed, and values less than 0.3 were also re-examined and corrected as suggested by Lowry and Gaskin (40), the software was configured for confirmatory factor analysis with the following setting:

- PLS Algorithm
- Weighting Scheme: Factor
- Maximum Iteration:1000
- Stop Criterion:7

After running the software with the aforementioned settings, the results of the confirmatory factor analysis were examined using *factor loading* which is visualized in the software as *Outer Loading* (Figure 3). when the factor loading of an observed variable exceeds 0.6, it is accepted, otherwise there is no significant relationship between this variable and the associated latent variable (44).

As shown in Figure 3 and According to Hair et al. (44), there is no significant relationship between the observed variables *Specialized Networks* and *Organizational Structure* with the "*Resources*" latent variable, having factor loadings of 0.492 and 0.391 (they are less than 0.6). Therefore, these two variables are removed from the model. But since the factor loadings of other relationships is greater than 0.6, it can be concluded that there is a significant relationship between the related observed and latent variables. for example, factor loading index between the observed variables *Readiness, Self-Organization* and *Flexibility* with the "*Adaptability*" latent variable, are 0.826, 0.892 and 0.804, respectively; Therefore, these variables are maintained in the model. This fact is also true in relation to other variables in the model.

The software is run again with the previous settings. Figure 4 shows the obtained results that is based on the updated factor load values. There is a significant relationship between all the observed and latent variables in the modified model. As shown in Figure 4, there is no factor loading less than 0.6. for example, factor loading index between the observed variables *Redundancy, Business Environments, Vulnerability Management* and *Governance Conditions* with the "*Vulnerability*" latent variable, are 0.643, 0.750, 0.892 and 0.843, respectively; Therefore, there is a significant relationship between all the mentioned observed and latent variables in this modified model.

After ensuring the establishment of a meaningful relationship between the observed variables and latent variables; it is essential to analyse their validity and

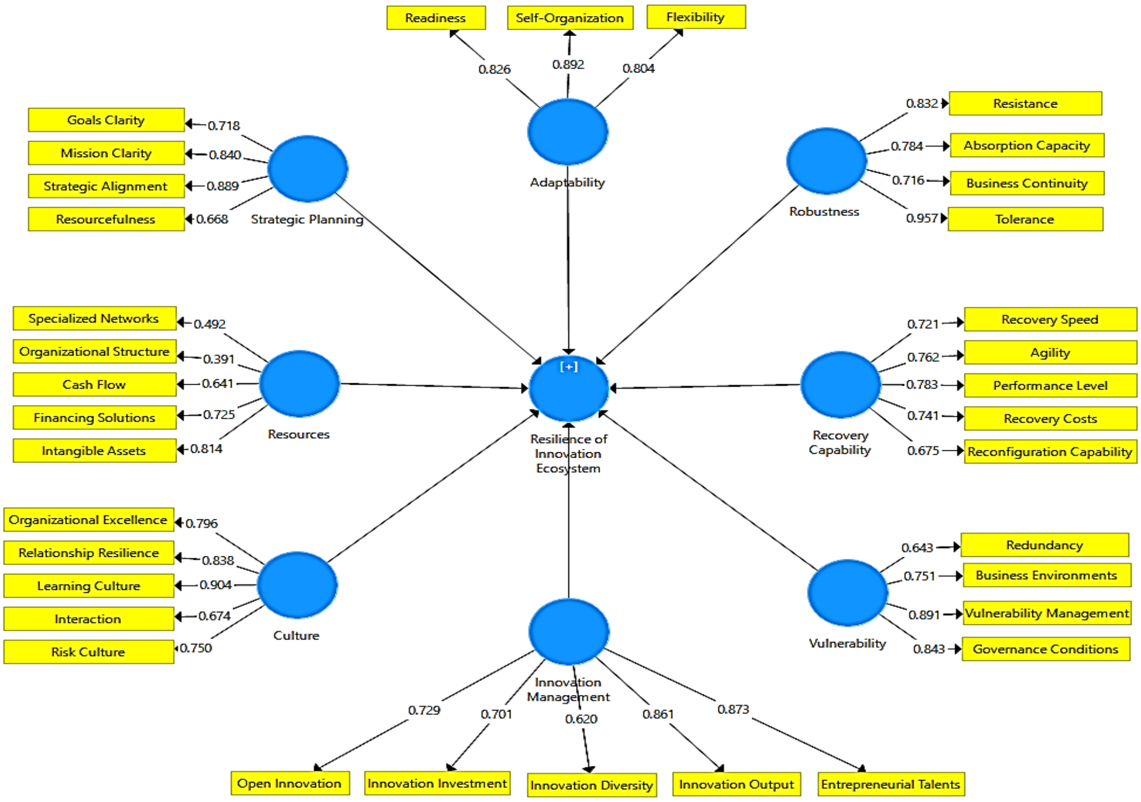


Figure 3. Factor loading values of observed variables and related latent variables in the primary model

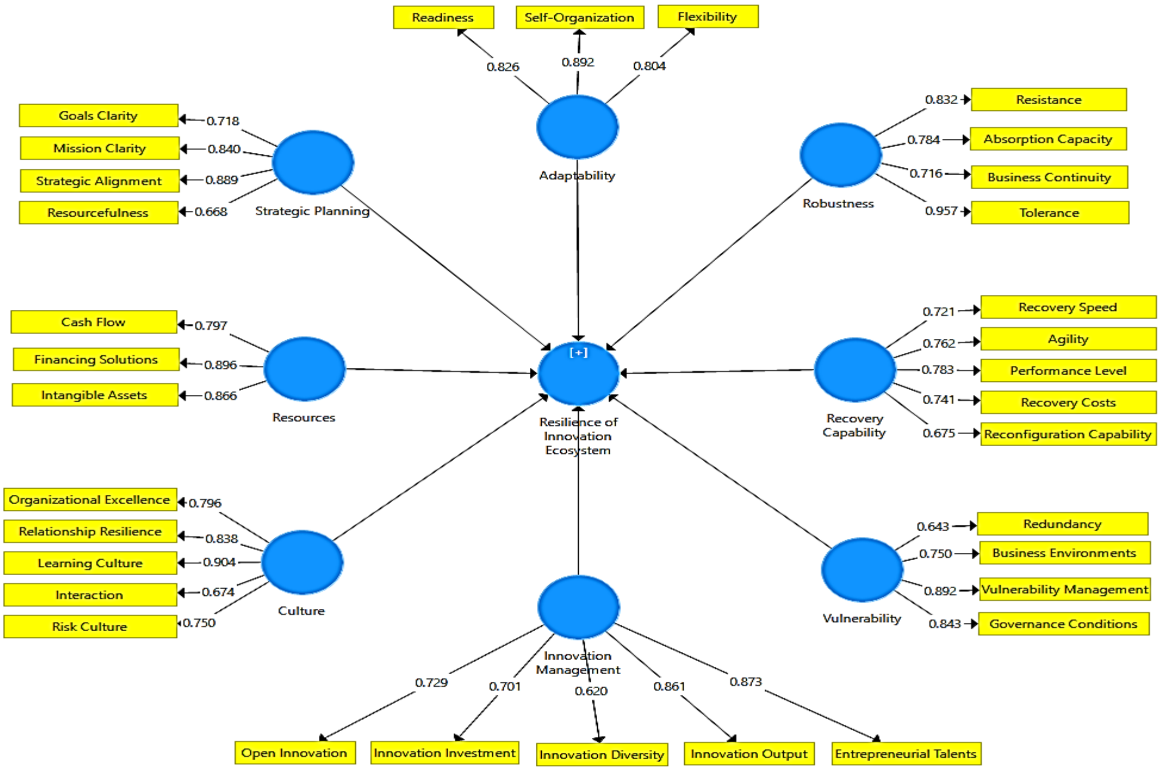


Figure 4. Factor loading values of observed variables and related latent variables in the modified model

reliability. To achieve this, several measures were calculated and analyzed, including *Cronbach's alpha*, *Composite Reliability (CR)*, *Factor loading*, and *Average Variance Extracted (AVE)*. Reliability analysis is conducted for the scales using Cronbach's alpha. Normally reliability coefficient of Cronbach's alpha ranges between 0 and 1. Greater or equal to 0.80 for a good scale, 0.70 for an acceptable scale and 0.60 for a scale for exploratory purposes (72). Composite reliability (CR) is a preferred alternative to Cronbach's alpha as a test of convergent validity in a reflective model. Composite reliability varies from 0 to 1, with 1 being perfect estimated reliability. Average variance extracted (AVE) may be used as a test of both convergent and divergent validity. AVE reflects the average communality for each latent factor in a reflective model. In general, According to previous researches, to create a model with appropriate validity and reliability, Cronbach's alpha value should be greater than 0.6 (73), and the value of the CR index should also be greater than 0.6 (47), in addition to the mentioned indicators, the AVE should also be greater than 0.5 (74) correspondingly the value of rho_A should be greater than 0.7 (75). As mentioned before, factor loading value should be more than 0.6. The aforementioned indicators are shown in Table 3; which shows that the framework of the factors affecting the resilience of the innovation ecosystem along with the relevant questions have appropriate validity and reliability.

TABLE 3. Validity and reliability indices associated with the modified framework after CFA

Construct	Item	Factor Loading	Cronbach's Alpha	Rho_A	Composite Reliability (Cr)	Average Variance Extracted (Ave)
Adaptability	Readiness	0.826	0.796	0.835	0.879	0.708
	Self-Organization	0.892				
	Flexibility	0.804				
Innovation Management	Open Innovation	0.729	0.824	0.878	0.873	0.582
	Innovation Investment	0.701				
	Innovation Diversity	0.620				
	Innovation Output	0.861				
	Entrepreneurial Talents	0.873				
Recovery Capability	Recovery Speed	0.721	0.796	0.811	0.856	0.544
	Agility	0.762				
	Performance Level	0.783				
	Recovery Costs	0.741				
	Reconfiguration Capability	0.675				
Culture	Organizational Excellence	0.796	0.854	0.876	0.896	0.634
	Relationship Resilience	0.838				
	Learning Culture	0.904				
	Interaction	0.674				
	Risk Culture	0.750				
Resource	Cash Flow	0.797	0.823	0.888	0.89	0.73
	Financing Solutions	0.896				
	Intangible Assets	0.866				
Robustness	Resistance	0.832	0.841	0.862	0.895	0.684
	Absorption Capacity	0.784				
	Business Continuity	0.716				
	Tolerance	0.957				
Strategic Planning	Goals Clarity	0.718	0.784	0.798	0.863	0.615
	Mission Clarity	0.840				
	Strategic Alignment	0.889				
	Resourcefulness	0.668				

Vulnerability	Redundancy	0.643	0.804	0.868	0.866	0.621
	Business Environments	0.750				
	Vulnerability Management	0.892				
	Governance Conditions	0.843				

4. 5. Structural Model Analysis

After the confirmatory factor analysis, we investigated the causal relationship between the independent and dependent variables, as well as investigating the Hypothesis raised in the context of the relationship between the independent variables and the dependent variable of innovation ecosystem resilience in the primary framework. Then, based on confirmed and meaningful Hypothesized path, a comprehensive framework of factors affecting the resilience of the innovation ecosystem was formed. The Hypothesis:

- H1: The resilience of innovation ecosystem is influenced by the adaptability factor (AD ⇨ RE)
- H2: The resilience of innovation ecosystem is influenced by innovation management factor (IM⇨RE)
- H3: The resilience of innovation ecosystem is influenced by the recovery capability factor (RY⇨RE)
- H4: The resilience of innovation ecosystem is influenced by the culture factor (RC ⇨ RE)
- H5: The resilience of innovation ecosystem is influenced by the resources factor (RS ⇨ RE)
- H6: The resilience of innovation ecosystem is influenced by the robustness factor (RO ⇨ RE)
- H7: The resilience of innovation ecosystem is influenced by the strategic planning factor (SP ⇨ RE)
- H8: The resilience of innovation ecosystem is influenced by the vulnerability factor (VU ⇨ RE)

As stated, at this stage of structural model analysis, we ensured the significance of the relationships between the observed variables related to each latent variable. Additionally, we verified the validity and reliability of the entire model. Subsequently the questionnaire was sent to a large number of experts from the target community. A total of 185 correct questionnaires were collected from different actors within the innovation ecosystem, and the structural model analysis process was performed in *SmartPLS* software.

The respondents must be familiar with the power innovation ecosystem. They should participate in at least one of the innovative technological projects in power innovation ecosystem, or having an organization position in one of the players of the innovation ecosystem, as well as a history of performing activities related to the performance indicators of the power innovation ecosystem. The personal information of the respondents is in Table 4. The majority of respondents had a master's degree. The gender of most of them was male and they were between 35 and 50 years old.

TABLE 4. Descriptive statistics of personal information

No.	Category	Item	Frequency
1	Gender	Male	160
		Female	25
2	Age	Under 35	37
		35–50	106
		Over 50	42
3	Education level	BS	11
		MS	123
		PHD	51

The settings applied in software to analyse the structural model are as follows:

- Bootstrapping (complete)
- Subsample: 5000
- Weighting Scheme: path
- Maximum Iteration:1000
- Stop Criterion:7

Bootstrapping option has been used to determine the statistical significance of the path coefficient and to calculate the t-values in this study. In Table 5, for each Hypothesized path, the indicators of *path coefficient*, *standard deviation*, *t-statistic*, *p-value* and Final Result are presented. According to the values in Table 5, it can be concluded that all the Hypothesized path with the t-values above 1.96 ($\alpha = 0.05$; two-sided test) and p-value less than 0.05 are significant with a probability of about 95% and the Hypothesized path with the t-values above 2.58 ($\alpha = 0.01$; two-sided test) and p-value less than 0.01 are significant with a probability of about 99% (74, 75).

According to the stated conditions, the t-values of the hypothesized path of AD and RE is 3.964, which is above 1.96 ($\alpha = 0.05$; two-sided test) and p-value is 0. So, the hypotheses path of AD and RE of the inner model is statistically significant with a probability of about 95%. It means that resilience of innovation ecosystem is influenced by adaptability factor.

The t-values of the hypothesized path of IM and RE is 1.981, which is above 1.96 ($\alpha = 0.05$; two-sided test) and p-value is 0.033. So, the hypotheses path of IM and RE of the inner model is statistically significant. It means that resilience of innovation ecosystem is influenced by Innovation Management factor.

Another t-values of the hypothesized path of RY and RE is 2.062, which is above 1.96 (α =0.05; two-sided test) and p-value is 0.039. So, the hypotheses path of RY and RE of the inner model is statistically significant. It means that resilience of innovation ecosystem is influenced by Recovery Capability factor.

Another t-values of the hypothesized path of RC and RE is 2.101, which is above 1.96 (α =0.05; two-sided test) and p-value is 0.036. So, the hypotheses path of RC and RE of the inner model is statistically significant. It means that resilience of innovation ecosystem is influenced by Culture factor.

The t-values of the hypothesized path of RS and RE is 2.403, which is above 1.96 (α =0.05; two-sided test) and p-value is 0.016. So, the hypotheses path of RS and RE of the inner model is statistically significant. It means that resilience of innovation ecosystem is influenced by Resources factor.

Another t-values of the hypothesized path of RO and RE is 2.777, which is above 1.96 (α =0.05; two-sided test) and p-value is 0.005. So, the hypotheses path of RO and RE of the inner model is statistically significant. It means that resilience of innovation ecosystem is influenced by Robustness factor.

The t-values of the hypothesized path of SP and RE is 3.307, which is above 1.96 (α =0.05; two-sided test) and p-value is 0.001. So, the hypotheses path of SP and RE of the inner model is statistically significant. It means that resilience of innovation ecosystem is influenced by Strategic Planning factor.

The t-values of the hypothesized path of VU and RE is 9.518, which is above 1.96 (α =0.05; two-sided test) and p-value is 0. So, the hypotheses path of VU and RE of the inner model is statistically significant. It means that resilience of innovation ecosystem is influenced by vulnerability factor.

4. 6. Providing a Comprehensive Framework and Discussion

According to the results (α=0.05) of Table 5 and the confirmed Hypothesis, the comprehensive framework of factors affecting the resilience of the innovation ecosystem is shown in Figure 5.

TABLE 5. Analysis of the structural model and Result of Hypothesis

Hypothesized Path (Inner Model)	Coefficient Sample Mean	Standard Deviation	T-Statistics (O/STDEV)	P-Values	Results (α=0.01)	Results (α=0.05)
H1: (AD ⇨ RE)	0.258	0.068	3.964	0	supported	supported
H2: (IM ⇨ RE)	0.094	0.041	1.981	0.033	Not supported	supported
H3: (RY ⇨ RE)	0.165	0.073	2.062	0.039	Not supported	supported
H4: (RC ⇨ RE)	0.102	0.048	2.101	0.036	Not supported	supported
H5: (RS ⇨ RE)	0.125	0.052	2.403	0.016	Not supported	supported
H6: (RO ⇨ RE)	0.161	0.058	2.777	0.005	supported	supported
H7: (SP ⇨ RE)	0.262	0.083	3.307	0.001	supported	supported
H8: (VU ⇨ RE)	0.498	0.052	9.518	0	supported	supported

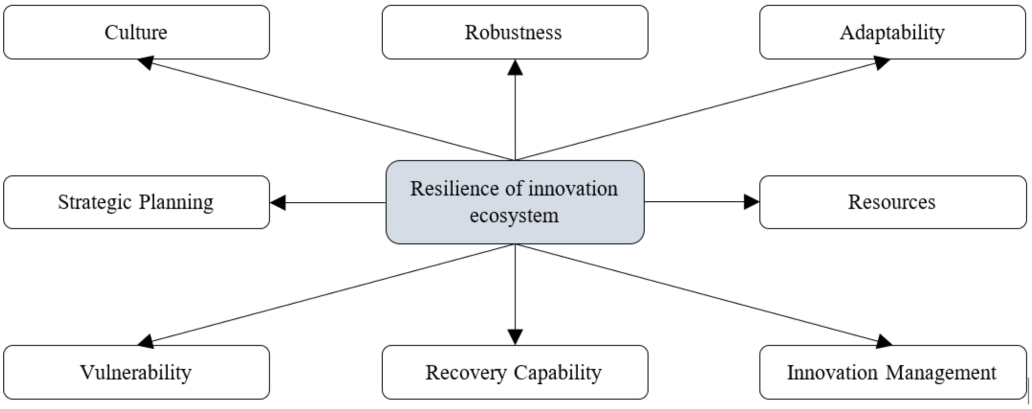


Figure 5. The comprehensive framework of factors affecting the resilience of the innovation ecosystem

Each factor of aforementioned framework is described as follows:

4. 2. 1. Culture as an Influent Factor on Innovation Ecosystem Resilience

The culture factor in a resilient innovative ecosystem refers to concepts such as Resilience thinking culture, Organizational Excellence, Relationship Resilience, Interaction, Risk Culture and Learning Culture. Resilience thinking culture among the actors of the innovation ecosystem has an important effect on the formation of the resilient innovation ecosystem. This culture includes resilience thinking at the individual level and at the organizational level in the entire ecosystem. Resilience thinking at the individual level refers to the spirit of solidarity and cooperation among employees, and the stronger this thinking is with the formation of resilient relationships and networks between stakeholders; The greater the flexibility of the organization and subsequently the entire ecosystem has increased so that the innovation ecosystem can unite in the face of crisis and create good strategic guidelines to deal with such situations (34). Resilience thinking along with raising awareness enables organizations to anticipate potential crises in the ecosystem to reduce the costs of crises when they occur (34). The culture of an organization shapes the interactive spirit of employees and their commitment to the organization. It encourages them to have a long-term commitment to the organization and they took appropriate steps towards organizational excellence. Risk management culture can be defined as the process of identifying, analyzing, monitoring and developing plan to response the risks to an acceptable level at an acceptable cost in different organization. In fact, risk management involves understanding the threats that can potentially affect performance and resilience. Organizational learning is one of the most important cultural components of a resilient organization, which helps the organization in gaining lessons learned from the failures of similar organizations, raises awareness in the direction of resilience and aids in planning for its improvement. Furthermore gives meaning to the cooperation of all employees of an organization, in this direction communication resilience helps organizations to create mutually beneficial relationships between different stakeholders enabling them to effectively address crises (27).

4. 2. 2. Strategic Planning as an Influent Factor on Innovation Ecosystem Resilience

Strategic resilience in an organization means achieving a balance between strategic stability (27) and flexibility in the organization's strategic plans to deal with various types of disruption, which requires that destructive events be predicted in the first step and its adverse consequences be limited (76). Clarity in Goals and Mission, Strategic Alignment and Resourcefulness are elements of Strategic resilience. That means Mission, vision and long-term, mid-term and short-term objectives should clearly define at different levels of organizations in innovation ecosystem. Strategic Alignment in ecosystem means that Strategic goals aligned with operational plans of different organization levels. Resourcefulness can be further considered as the ability to use resources (such as financial, physical, technological and informational) and human resources to meet set priorities and achieve goals. Anticipation of destructive events is one of the most important characteristics of resilience. This feature can be used to increase the adaptability and resistance of ecosystems by predicting and responding appropriately to malicious events. Anticipation is defined as the ability to learn, adapt, take preventive action, resist and recover from catastrophic events (18). After forecasting, the next crucial step is to prepare plans before the occurrence of destructive events in order to reduce vulnerability, prevent secondary disasters following the initial event, and establish short-term and long-term priorities after the disaster (21). To implement the prepared plans, it is essential to engage in effective planning, which involves setting priorities and mobilizing resources.

4. 2. 3. Resources as an Influent Factor on Innovation Ecosystem Resilience

Resources are key factors in creating innovation ecosystem resilience and determine the ability of ecosystem actors to improve resilience (34). Resources include wide range of elements such as human resources, financial resources, social capital, diversity of talent and personal skills, organizational structure, information systems, and intangible assets. The proper management of resources improves the supportive infrastructure and provides the necessary implementation of the strategic plans of an organization. Generally, in a resilient innovation ecosystem, organizations have a suitable cash flow and several financing solutions.

One of the most important resources in recent years is information and communication technology, this kind of resource has increased the awareness by creating a flow of valuable information from within and outside the ecosystem. As a result, it plays an effective role in identifying, monitoring and controlling risks and ultimately helps to improve the resilience of the organization. Organizational structure is also considered as another important resource. The more complex the internal structure of the innovation ecosystem and the greater its diversity, the stronger its internal support capacity and higher its external resistance ability; and this is due to increase in the proportion of innovation-related sectors as well as multiple aligned activities (6). In general, a strong and flexible organizational structure enables organizations to grow by opportunities and improve their overall resistance to risk (34).

4. 2. 4. Innovation Management as an Influent Factor on Innovation Ecosystem Resilience

Concepts such as Implementing Open Innovation, diversity of innovators and innovative subjects, Innovation Output and Appropriate Investments in innovative subjects refers to Innovation Management factor. Regarding to Open innovation concept organization in ecosystem does not just rely on their own internal knowledge and resources for innovation but also uses multiple external sources to drive innovation.

In natural ecosystems, biodiversity is a key factor that affects the resilience of the ecosystem. Similarly, in the innovation ecosystem, the diversity of innovators and innovative subjects is an important basis for establishing stable connections between various topics especially when faced with external shocks (6). In order to examine the diversity, flow and output of innovation in an innovation ecosystem, several indicators may be used. These may include: the number of registered patents, scientific publications including scientific articles and books, the proportion of researchers, the proportion of companies with dedicated research and development units, the number of research and development institutions, the number of products commercialized, the number of technology contracts in the technology market, etc. In an innovation ecosystem, greater diversity and innovation flow lead to enhanced adaptability to disturbances and by implementing proper planning, the resilience of the ecosystem against risks and disturbances can also be boosted.

4. 2. 5. Vulnerability as an Influent Factor on Innovation Ecosystem Resilience

Vulnerability is the measure of an ecosystem's susceptibility to damage caused by a disruption. It is an undesirable concept, as it highlights the potential weaknesses in an ecosystem.

Redundancy, as one of the elements of evaluating vulnerability, refers to the degree to which player, or other units of an innovation ecosystem are replaceable and can meet functional needs in the event of destruction. Also, the stability of the business environment (economic, political and social conditions) and governance conditions are other elements that affect vulnerability. Vulnerability assessment has been developed in two areas of natural hazards and social science communities (5). In the field of natural hazards, vulnerability is defined as a combination of risk factors and the likelihood of ecosystem losses. In the field of social science societies, vulnerability focuses on inequality of sensitivity and exposure to threat (social equity). Vulnerability assessment is often used as a preventive planning tool or post-event analysis. Therefore, to enhance the innovation ecosystem's resilience, we must first identify and monitor vulnerable points. Then, we need to create and implement a timely plan to address any disruptions. One of the most influential factors influencing the vulnerability of actors in the innovation ecosystem is the governance, social, economic and political conditions, that govern the business environment. For example, the fluctuations of the economic cycle affect the political environment and the market environment. Therefore, it will have a direct impact on the business of innovation ecosystem actors (34).

4. 2. 6. Robustness as an Influent Factor on Innovation Ecosystem

Robustness is defined as the ability of ecosystems' elements to withstand a certain level of disturbance without experiencing major change in their performance (14). A resilient innovation ecosystem must be resistant to uncertainties and disruptions and increase its tolerance level. Tolerance represents the persistence of the innovation ecosystem in uncertain situations. Ecosystem with a high level of tolerance have a positive attitude towards challenges and see them as an opportunity for ideation and innovative behaviour. Therefore, different players within the innovation ecosystem should improve absorption capacity of their organization which is defined as a measure of an ecosystem's ability to resist turbulent conditions and reduce consequences; make appropriate strategic planning for its improvement. We can outline the capacity to absorb and the level of tolerance and resistance of an organization expressed in robustness.

4. 2. 7. Adaptability as an Influent Factor on Innovation Ecosystem Resilience

In general, Adaptability includes concepts such as Readiness, Flexibility and Self-Organization. Readiness means the innovation ecosystem can well prepared to respond to changes and disruptions. Flexibility in an innovation ecosystem causes to its players and their operational processes to have appropriate flexibility when it faces disruption and change. Self-Organization refers to ability of innovation ecosystem and its players to adapt and organize itself after disruptions. In this regard, adaptive capacity is very important. Adaptive capacity is defined as the extent to which an ecosystem can organize itself and overcome a disruptive event to restore ecosystem performance and overcome the disturbance without requiring any recovery activities. The knowledge learned through disruptive event, system reconfiguration, and trained personnel are the most adaptive activities that contribute to resilience (18). Actually, The concept of adaptive capacity is known as the ability for responding to disruption (5). Organizations in an innovation ecosystem need to focus on improving transformative capacity and evolutionary capacity in addition to adaptive capacity. Evolutionary capacity refers to the ability of an innovation ecosystem to gradually evolve and develop over time. Evolutionary capacity is built on adaptability and its purpose is to change or restore the environment in which the innovation ecosystem is set, so that the organization can develop in a better direction. Only by continuously evolving in a complex environment, the innovation ecosystem can resist environmental uncertainties and move towards sustainable development. A stronger evolutionary capacity results in a greater resilience value for the innovation ecosystem, allowing organizations to thrive and survive. On the contrary, if we face weaker evolutionary capacity, the adaptation capacity of the innovation ecosystem decreases; and a lower resilience value leads to weaker the organization's competitiveness and it encounter the risk of being vanished from the market (34).

4. 2. 8. Recovery Capability as an Influent Factor on Innovation Ecosystem Resilience The characteristic of a resilient ecosystem is its recovery and restoration ability after a destructive event. Concepts such as Recovery Speed, Agility, Performance Level, Recovery Costs and Reconfiguration Capability constitute the Recovery Capability factor in an innovation ecosystem. Different players of innovation ecosystem must have appropriate agility in order to deal with ecosystem change. It is necessary to focus on quickly recover and return to desired performance level in case of ecosystem disruption. In this regard, recovery speed and recovery cost are imperative components.

5. CONCLUSION AND FUTURE RESEARCH

Due to environmental uncertainties as well as the high speed of changes, organizations face various and intermittent threats. Therefore, the innovation ecosystem as a network of interconnected organizations such as universities, knowledge-based companies, etc. always faces destructive events such as international commercial limitations, competitive pressures, loss of experts, technology changes, financial crises, etc. Improving the resilience of this kind of ecosystems means enhancing the ability of its various actors to absorb, adapt and recover from these crises as well as increasing the assurance of maintaining its performance at the desired level; In return, these efforts could result in the sustainable development of countries. As stated before, we find that a holistic view of innovation ecosystem resilience has received less attention and so far, we do not explore any comprehensive framework of factors affecting resilience. Therefore, in this article, in response to the research question, we investigate what factors affect the resilience of the innovation ecosystem in a holistic perspective, and what are the connections between them and resilience.

First, based on the meta-Synthesis method, we systematically reviewed past related research, extracted the concept and analysed them. After validation of findings (based on content analysis method) we presented a primary framework of factors affecting the innovation ecosystem resilience. Then, in order to develop a comprehensive framework of affecting factors on resilience of the innovation ecosystem, Confirmatory factor analysis and structural equation modelling (SEM) were used. In the process of CFA, we identified 39 observed variables and used them to create a questionnaire then analysed information gathered from the target population. After removing any unrelated variables, in the second step, the structural model analysis was done and eight main hypothesized paths were tested to investigate the causal relationships between the identified factors (independent variables) and the dependent variable under the title of "resilience of the innovation ecosystem". Subsequently the comprehensive framework of affecting factors on resilience of the innovation ecosystem was presented. The results indicate that the factors of vulnerability and adaptability are the most important influencing ones on the resilience of the innovation ecosystem, and the recovery capacity and culture of resilience factors are less influential than other independent variables. In addition to the innovation of using the meta-synthesis qualitative method in order to provide the initial framework, finalizing the effective factors using the structural equation method, and providing a comprehensive framework of the effective factors on resilience of the innovation ecosystem, are also other novelties of this article. On the other hand, due to the diversity of players, the architecture and internal cooperation model modification of different innovation ecosystems, we benefited the experts' opinions of the power innovation ecosystem affiliated to Iranian power industry (as a case study), besides we tailored and coordinated the resilience assessment framework for the Iranian power innovation ecosystem, which are another witness for the uniqueness of this research.

According to the presented framework, the leaders and coordinators of the innovation ecosystem should emphasize several concepts, including the promotion of resilience thinking throughout the entire ecosystem. This includes resilience thinking at the individual level and the organizational level throughout innovation ecosystem; and it can shape the foundation of appropriate relationships and networking between stakeholders, improv the interactive spirit, increase commitment to the goals of the entire ecosystem, and finally making effective alliance against disruptions. Identifying vulnerable points and predicting destructive events are also among the most important features of resilience that should be considered by managers. This feature can be used to increase adaptability and robustness by appropriate anticipating and responding to disruptive events. In addition, forecasting, formulating and implementing strategic plans are also useful means to deal with all kinds of disturbances. Obviously, the proper management of resources (including human resources, financial resources, social capital, diversity of talent and personal skills, organizational structure, information systems, and etc.) improves the supportive infrastructures along with implementation of an organization's strategic plan. In addition, the more innovation policymakers support open innovation events, diversity of the innovation flow in an innovation ecosystem would become stronger and more stable; so, the ecosystem can easily adapt to disruptions. In order to create a resilient innovation ecosystem, different actors of the ecosystem also need to focus on the adaptive capacity and improve the transformative and evolutionary capacity.

The results of this research help the policy makers of technology development and innovation to evaluate the resilience of their governed innovation ecosystems based on the mentioned framework, and with the aim of achieving intended objectives, they would make appropriate strategies considering effective factors by identifying the strengths and weaknesses; and improve the ecosystem resilience. Considering the importance of innovation in improving the national economy, the realization of sustainable development of societies seems to be more accessible in practice by being more resilient in innovation ecosystem.

Due to variety of conditions governing the internal and external environment of innovation ecosystems, it is suggested that future research explores and analyzes the proposed framework in this article across various statistical communities, industries, and countries.

6. REFERENCES

1. Yin D, Ming X, Zhang X. Sustainable and smart product innovation ecosystem: An integrative status review and future perspectives. Journal of Cleaner Production. 2020;274:123005. 10.1016/j.jclepro.2020.123005
2. Jucevičius G, Grumadaitė K. Smart development of innovation ecosystem. Procedia-social and behavioral sciences. 2014;156:125-9. 10.1016/j.sbspro.2014.11.133
3. Engler J, Kusiak A. Modeling an innovation ecosystem with adaptive agents. International journal of innovation science. 2011;3(2):55-68. 10.1260/1757-2223.3.2.55
4. Robertson J, Caruana A, Ferreira C. Innovation performance: The effect of knowledge-based dynamic capabilities in cross-country innovation ecosystems. International Business Review. 2021:101866. 10.1016/j.ibusrev.2021.101866
5. Gillespie-Marthaler L, Nelson KS, Baroud H, Kosson DS, Abkowitz M. An integrative approach to conceptualizing sustainable resilience. Sustainable and Resilient Infrastructure. 2019;4(2):66-81. 10.1080/23789689.2018.1497880
6. Bai T, Li J. Resilience Evaluation of Innovation Ecosystem of High-tech Enterprises in Liaoning Province. Financial Engineering and Risk Management. 2022;5(1):61-8. 10.23977/ferm.2022.050110
7. Hosseini S, Barker K, Ramirez-Marquez JE. A review of definitions and measures of system resilience. Reliability Engineering & System Safety. 2016;145(1):47-61. 10.1016/j.ress.2015.08.006
8. Nelson K, Gillespie-Marthaler L, Baroud H, Abkowitz M, Kosson D. An integrated and dynamic framework for assessing sustainable resilience in complex adaptive systems. Sustainable and Resilient Infrastructure. 2020;5(5):311-29. 10.1080/23789689.2019.1578165
9. Nylund PA, Ferras-Hernandez X, Brem A. Strategies for activating innovation ecosystems: introduction of a taxonomy. IEEE Engineering Management Review. 2019;47(4):60-6. 10.1109/EMR.2019.2931696
10. Xiao L, Cao H, editors. Organizational resilience: The theoretical model and research implication. ITM Web of Conferences; 2017: EDP Sciences. 10.1051/itmconf/20171204021
11. Walker B, Holling CS, Carpenter SR, Kinzig A. Resilience, adaptability and transformability in social–ecological systems. Ecology and society. 2004;9(2). 10.5751/ES-00650-090205
12. Herrera H, Kopainsky B. Using system dynamics to support a participatory assessment of resilience. Environment systems and decisions. 2020;40(3):342-55. 10.1007/s10669-020-09760-5
13. Egli L, Weise H, Radchuk V, Seppelt R, Grimm V. Exploring resilience with agent-based models: state of the art, knowledge gaps and recommendations for coping with multidimensionality. Ecological Complexity. 2019;40(2):100718. 10.1016/j.ecocom.2018.06.008
14. Bruneau M, Chang SE, Eguchi RT, Lee GC, O'Rourke TD, Reinhorn AM, et al. A framework to quantitatively assess and enhance the seismic resilience of communities. Earthquake spectra. 2003;19(4):733-52. 10.1193/1.1623497
15. Sauser B, Baldwin C, Pourreza S, Randall W, Nowicki D. Resilience of small-and medium-sized enterprises as a correlation to community impact: an agent-based modeling approach. Natural Hazards. 2018;90(1):79-99. 10.1007/s11069-017-3034-9
16. Henry D, Ramirez-Marquez JE. Generic metrics and quantitative approaches for system resilience as a function of time. Reliability Engineering & System Safety. 2012;99(1):114-22. 10.1016/j.ress.2011.09.002
17. Sweetapple C, Fu G, Butler D. Reliable, robust, and resilient system design framework with application to wastewater-treatment plant control. Journal of Environmental Engineering. 2017;143(3):04016086. 10.1061/(ASCE)EE.1943-7870.0001171
18. Ahmadi S, Saboohi Y, Vakili A. Frameworks, quantitative indicators, characters, and modeling approaches to analysis of energy system resilience: A review. Renewable and Sustainable Energy Reviews. 2021;144(10):110988. 10.1016/j.rser.2021.110988
19. Tran HT, Balchanos M, Domerçant JC, Mavris DN. A framework for the quantitative assessment of performance-based system resilience. Reliability Engineering & System Safety. 2017;158:73-84. 10.1016/j.ress.2016.10.014
20. Biddle L, Wahedi K, Bozorgmehr K. Health system resilience: a literature review of empirical research. Health policy and planning. 2020;35(8):1084-109. 10.1093/heapol/czaa032
21. Zhu J, Manandhar B, Truong J, Ganapati NE, Pradhananga N, Davidson RA, et al. Assessment of infrastructure resilience in the 2015 Gorkha, Nepal, earthquake. Earthquake Spectra. 2017;33(1_suppl):147-65. 10.1193/121116eqs231m
22. Yu S, Kim S-W, Oh C-W, An H, Kim J-M. Quantitative assessment of disaster resilience: An empirical study on the importance of post-disaster recovery costs. Reliability Engineering & System Safety. 2015;137:6-17. 10.1016/j.ress.2014.12.007
23. Burnard KJ, Bhamra R. Challenges for organisational resilience. Continuity & Resilience Review. 2019. 10.1108/CRR-01-2019-0008
24. Hillmann J, Guenther E. Organizational resilience: a valuable construct for management research? International Journal of Management Reviews. 2021;23(1):7-44. 10.1111/ijmr.12239

25. Ma Z, Xiao L, Yin J. Toward a dynamic model of organizational resilience. Nankai Business Review International. 2018;9(1):246–63. 10.1108/NBRI-07-2017-0041

26. Hamel G, Valikangas L. The quest for resilience, Harvard business review, September, 2003. 2003.

27. Chen R, Xie Y, Liu Y. Defining, conceptualizing, and measuring organizational resilience: A multiple case study. Sustainability. 2021;13(5):2517. 10.3390/su13052517

28. Annarelli A, Battistella C, Nonino F. A framework to evaluate the effects of organizational resilience on service quality. Sustainability. 2020;12(3):958-73. 10.3390/su12030958

29. Lee AV, Vargo J, Seville E. Developing a tool to measure and compare organizations' resilience. Natural hazards review. 2013;14(1):29-41. 10.3390/su12030958

30. McManus S, Seville E, Vargo J, Brunsdon D. Facilitated process for improving organizational resilience. Natural hazards review. 2008;9(2):81-90. 10.1061/(ASCE)1527-6988(2008)9:2(81)

31. Rahi K. Indicators to assess organizational resilience–a review of empirical literature. International Journal of Disaster Resilience in the Built Environment. 2019;10(2/3):85-98. 10.1108/IJDRBE-11-2018-0046

32. Sanchis R, Poler R. Enterprise resilience assessment—A quantitative approach. Sustainability. 2019;11(16):4327. 10.3390/su11164327

33. Aleksić A, Stefanović M, Arsovski S, Tadić D. An assessment of organizational resilience potential in SMEs of the process industry, a fuzzy approach. Journal of Loss Prevention in the Process Industries. 2013;26(6):1238-45. 10.1016/j.jlp.2013.06.004

34. Tong Y, Liu Q, He K, Liu M. Research on the Influencing Factors of Innovation Ecosystem Resilience of High-tech Enterprises. Industrial Engineering and Innovation Management. 2022;5(1):57-63. 10.23977/ieim.2022.050110

35. Roundy PT, Brockman BK, Bradshaw M. The resilience of entrepreneurial ecosystems. Journal of Business Venturing Insights. 2017;8:99-104. 10.1016/j.jbvi.2017.08.00

36. Sant TD, de Souza Bermejo PH, Moreira MF, de Souza WVB. The structure of an innovation ecosystem: foundations for future research. Management Decision. 2020;58(12):2725-42. 10.1108/MD-03-2019-0383

37. Sandelowski M, Barroso J. Handbook for synthesizing qualitative research: springer publishing company; 2006.

38. Bak O, Shaw S, Colicchia C, Kumar V. A systematic literature review of supply chain resilience in small–medium enterprises (SMEs): A call for further research. IEEE Transactions on Engineering Management. 2020;70(1):328-41. 10.1109/TEM.2020.3016988

39. Pal R, Torstensson H, Mattila H. Antecedents of organizational resilience in economic crises—an empirical study of Swedish textile and clothing SMEs. International Journal of Production Economics. 2014;147:410-28. 10.1016/j.ijpe.2013.02.031

40. Majnouni-Toutakhane A, Sareban VH. Promotion of Urban Resilience with Citizens' Local Participation Approach Case Study: Bonab City. Journal of Engineering Research. 2019;7(1).

41. Christopher M, Peck H. Building the resilient supply chain. 2004.

42. Petit F, Bassett G, Black R, Buehring W, Collins M, Dickinson D, et al. Resilience measurement index: An indicator of critical infrastructure resilience. Argonne National Lab.(ANL), Argonne, IL (United States); 2013.

43. Saad MH, Hagelaar G, van der Velde G, Omta S. Conceptualization of SMEs' business resilience: A systematic literature review. Cogent Business & Management. 2021;8(1):1938347. 10.1080/23311975.2021.1938347

44. Li T, Dong Y, Liu Z. A review of social-ecological system resilience: Mechanism, assessment and management. Science of the Total Environment. 2020;723:138113. 10.1016/j.scitotenv.2020.138113

45. Shirali GA, Mohammadfam I, Ebrahimipour V. A new method for quantitative assessment of resilience engineering by PCA and NT approach: A case study in a process industry. Reliability Engineering & System Safety. 2013;119:88-94.

46. Azadeh A, Salehi V, Ashjari B, Saberi M. Performance evaluation of integrated resilience engineering factors by data envelopment analysis: The case of a petrochemical plant. Process Safety and Environmental Protection. 2014;92(3):231-41. 10.1016/j.psep.2013.03.002

47. Jalali G, Tavakkoli-Moghaddam R, Ghomi-Avili M, Jabbarzadeh A. A network design model for a resilient closed-loop supply chain with lateral transshipment. International Journal of Engineering, Transactions C: Aspects. 2017;30(3):374-83. 10.5829/idosi.ije.2017.30.03c.07

48. Pant R, Barker K, Zobel CW. Static and dynamic metrics of economic resilience for interdependent infrastructure and industry sectors. Reliability Engineering & System Safety. 2014;125:92-102. 10.1016/j.ress.2013.09.007

49. Cutter SL, Barnes L, Berry M, Burton C, Evans E, Tate E, et al. A place-based model for understanding community resilience to natural disasters. Global environmental change. 2008;18(4):598-606. 10.1016/j.gloenvcha.2008.07.013

50. Rose AZ. Economic resilience to disasters. 2009(1).

51. Buliga O, Scheiner CW, Voigt K-I. Business model innovation and organizational resilience: towards an integrated conceptual framework. Journal of Business Economics. 2016;86(6):647-70. 10.1007/s11573-015-0796-y

52. Muller G. Fuzzy architecture assessment for critical infrastructure resilience. Procedia Computer Science. 2012;12:367-72. 10.1016/j.procs.2012.09.086

53. Rezaei M, Tolou-Askari M, Amirahmadi M, Ghods V. Challenges of Generation and Transmission Expansion Planning Considering Power System Resilience and Provide Solutions. International Journal of Engineering, Transaction B: Applications. 2023;36(5):824-41. 10.5829/IJE.2023.36.05B.01

54. Ates A, Bititci U. Change process: a key enabler for building resilient SMEs. International Journal of Production Research. 2011;49(18):5601-18. 10.1080/00207543.2011.563825

55. Biggs D, Hicks CC, Cinner JE, Hall CM. Marine tourism in the face of global change: The resilience of enterprises to crises in Thailand and Australia. Ocean & coastal management. 2015;105:65-74. 10.1016/j.ocecoaman.2014.12.019

56. Gunasekaran A, Rai BK, Griffin M. Resilience and competitiveness of small and medium size enterprises: an empirical research. International journal of production research. 2011;49(18):5489-509. 10.1016/j.ocecoaman.2014.12.019

57. Biggs D. Understanding resilience in a vulnerable industry: the case of reef tourism in Australia. Ecology and society. 2011;16(1). 10.5751/ES-03948-160130

58. Vlacheas P, Stavroulaki V, Demestichas P, Cadzow S, Ikonomou D, Gorniak S. Towards end-to-end network resilience. International Journal of Critical Infrastructure Protection. 2013;6(3-4):159-78. 10.1016/j.ijcip.2013.08.004

59. Henry D, Ramirez-Marquez JE. Generic metrics and quantitative approaches for system resilience as a function of time. Reliability Engineering & System Safety. 2012;99:114-22. 10.1016/j.ress.2011.09.002

60. Kahan JH, Allen AC, George JK. An operational framework for resilience. Journal of Homeland Security and Emergency Management. 2009;6(1). 10.2202/1547-7355.1675

61. Keck M, Sakdapolrak P. What is social resilience? Lessons learned and ways forward. Erdkunde. 2013:5-19. 10.3112/erdkunde.2013.01.02

62. Spiegler VL, Naim MM, Wikner J. A control engineering approach to the assessment of supply chain resilience. International journal of production research. 2012;50(21):6162-87. 10.1080/00207543.2012.710764

63. Percoco M. Infrastructure and economic efficiency in Italian regions. Networks and Spatial Economics. 2004;4(4):361-78. 10.1007/s11067-014-9261-7

64. Pettit TJ, Fiksel J, Croxton KL. Ensuring supply chain resilience: development of a conceptual framework. Journal of business logistics. 2010;31(1):1-21. 10.1002/j.2158-1592.2010.tb00125.x

65. Rose A. Economic resilience to natural and man-made disasters: Multidisciplinary origins and contextual dimensions. Environmental Hazards. 2007;7(4):383-98. 10.1016/j.envhaz.2007.10.001

66. Chang SE, Shinozuka M. Measuring improvements in the disaster resilience of communities. Earthquake spectra. 2004;20(3):739-55. 10.1193/1.1775796

67. Faturechi R, Levenberg E, Miller-Hooks E. Evaluating and optimizing resilience of airport pavement networks. Computers & Operations Research. 2014;43:335-48. 10.1016/j.cor.2013.10.009

68. Lawshe CH. A quantitative approach to content validity. Personnel psychology. 1975;28(4):563-75. 10.1111/j.1744-6570.1975.tb01393.x

69. Lowry PB, Gaskin J. Partial least squares (PLS) structural equation modeling (SEM) for building and testing behavioral causal theory: When to choose it and how to use it. IEEE transactions on professional communication. 2014;57(2):123-46. 10.1109/TPC.2014.2312452

70. Hair JF, Black WC, Babin BJ, Anderson RE, Tatham RL. Multivariate data analysis (Vol. 6): Pearson Prentice Hall Upper Saddle River. NJ; 2006.

71. Rouf MA, Akhtaruddin M. Factors affecting the voluntary disclosure: a study by using smart PLS-SEM approach. International Journal of Law and Management. 2018;60(6):1498-508. 10.1108/IJLMA-01-2018-0011

72. Taber KS. The use of Cronbach’s alpha when developing and reporting research instruments in science education. Research in science education. 2018;48:1273-96. 10.1007/s11165-016-9602-2

73. Bagozzi RP, Yi Y. On the evaluation of structural equation models. Journal of the academy of marketing science. 1988;16:74-94. 10.1007/BF02723327

74. Kline RB. Principles and practice of structural equation modeling: Guilford publications; 2023.

75. Dijkstra TK, Henseler J. Consistent partial least squares path modeling. MIS quarterly. 2015;39(2):297-316. 10.25300/MISQ/2015/39.2.02

76. Adger WN. Social and ecological resilience: are they related? Progress in human geography. 2000;24(3):347-64. 10.1191/030913200701540465

Persian Abstract

چکیده

امروزه اکوسیستم نوآوری به عنوان موتور محرک اقتصاد دانش بنیان کشورها درنظر گرفته می شود و تاب آوری آن در برابر اختلالات و حفظ عملکرد آن، از دغدغه بسیاری از سیاست‌گذاران می باشد. بنابراین یافتن چارچوب جامعی از فاکتورهای موثر بر تاب‌آوری اکوسیستم نوآوری می‌تواند در این راستا راهگشا باشد که این پژوهش به دنبال ارائه چنین چارچوب جامعی با استفاده از رویکرد فراترکیب، تحلیل عاملی تأییدی و مدل‌سازی معادلات ساختاری است و اکوسیستم نوآوری وزارت نیروی ایران نیز به عنوان مطالعه موردی در نظر گرفته شده است. فاکتورهای چارچوب فوق الذکر عبارتند از انطباق پذیری، مدیریت نوآوری، توانمندی بازیابی، فرهنگ، منابع، استحکام و پایداری، برنامه‌ریزی راهبردی و آسیب‌پذیری. نتایج نشان می‌دهد که فاکتورهای آسیب‌پذیری و انطباق پذیری مهم‌ترین فاکتورهای تاثیرگذار بر تاب‌آوری اکوسیستم نوآوری و فاکتورهای ظرفیت بازیابی و فرهنگ تاب‌آوری کمترین تاثیر را در این ارتباط دارند. استفاده از روش فراترکیب برای مرور نظام‌مند ادبیات، تحلیل محتوا و دسته‌بندی فاکتورهای موثر و همچنین ارائه یک چارچوب جامع از آنها مبتنی بر تحلیل عاملی و مدل‌سازی معادلات ساختاری از نوآوری‌های این پژوهش است. نتایج این پژوهش به سیاستگذاران اکوسیستم نوآوری کمک می‌نماید تا مبتنی بر چارچوب مذکور، به ارزیابی فاکتورهای مختلف پرداخته و سپس جهت تحقق اهداف مدنظر، برنامه‌ریزی راهبردی و عملیاتی نمایند.

IJE TRANSACTIONS B: Applications Vol. 37 No. 02, (February 2024) 341-351

International Journal of Engineering

Journal Homepage: www.ije.ir

A Distributed Cooperative Secondary Control Scheme for Obtaining Power and Voltage References of Distributed Generations in Islanded DC Microgrids

M. Nabian Dehaghani[a], M. Biglarahmadi[a], S. Y. Mousazadeh Mousavi*[b], M. Abdolahi[c]

[a] *Department of Electrical Engineering, University of Kashan, Kashan, Iran*
[b] *Department of Electrical Engineering, Faculty of Technology and Engineering, University of Mazandaran, Babolsar, Iran*
[c] *Department of Electrical Engineering, Faculty of Electrical and Computer Engineering, Babol Noshirvani University of Technology, Babol, Iran*

PAPER INFO

Paper history:
Received 18 August 2023
Received in revised form 14 September 2023
Accepted 16 September 2023

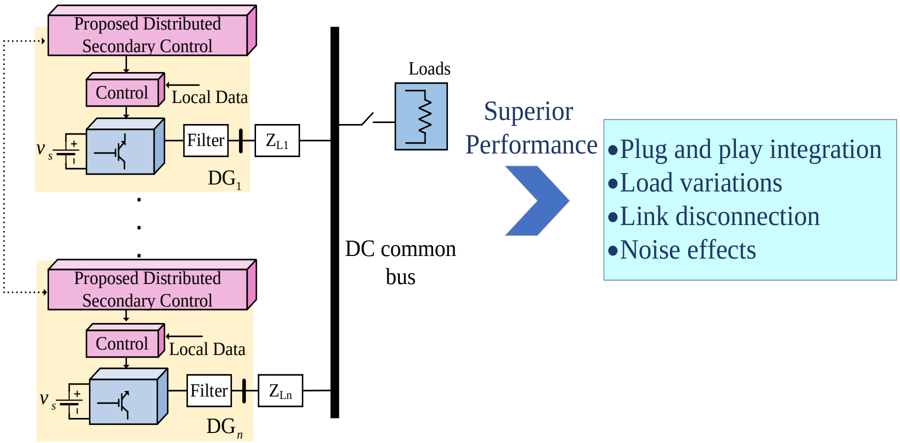
Keywords:
DC Microgrids
Distributed Secondary Control
Droop Control
Power Sharing
Voltage Regulation

ABSTRACT

Due to the widespread adoption of DC source-based distributed energy resources (DERs) and loads, alongside advancements in power electronics technology, DC microgrids (DC MGs) have recently gained significant attention. To effectively implement DC MGs, it is crucial to employ a suitable control strategy that maintains the bus voltage at the desired level and ensures appropriate power sharing among the integrated DERs. To address these objectives, a two-layer control scheme is proposed in this paper. In the primary layer, a customized droop control scheme is introduced, which applies lower voltage drop in comparison to the conventional droop strategies. Simultaneously, in the secondary layer, a modified voltage controller which is supplemented by a term to enhance power sharing in a distributed manner is employed,. The proposed control strategies are characterized by their simplicity and low communication infrastructure requirements. To assess the efficacy of the proposed control architecture, several case studies, including plug and play integration, load variations, and communication challenges, including link disconnection and noise effects, are conducted. Additionally, the performance of the proposed strategies is benchmarked against an architecture featuring conventional primary droop control and cooperative distributed secondary control approaches. The simulation studies conducted in MATLAB/SIMULINK software demonstrate that the proposed control methods outperform the alternative approaches, confirming their effectiveness in maintaining voltage regulation and power sharing objectives.

doi: *10.5829/ije.2024.37.02b.10*

Graphical Abstract



---

*Corresponding Author Email: s.y.mousazadeh@umz.ac.ir (S. Y. Mousazadeh Mousavi)

Please cite this article as: Nabian Dehaghani M, Biglarahmadi M, Mousazadeh Mousavi SY, Abdolahi M. A Distributed Cooperative Secondary Control Scheme for Obtaining Power and Voltage References of Distributed Generations in Islanded DC Microgrids, International Journal of Engineering, Transactions B: Applications. 2024;37(02):341-51.

1. INTRODUCTION

In recent years, there has been a growing interest in the microgrid (MG) concept as a means of reliability enhancement, improvement in power quality, and integration of renewable energy resources (RESs) (1, 2). The control of MGs, all types of AC, DC, and AC/DC MGs, in both their operational inslanded and grid-connected modes is the main interest of many researchers (3). The use of DC MGs has particularly gained attention due to the evolution in power electronics technology, more compatibility with other equipment, and greater flexibility in integrating of RERs and energy storage systems into the grid (4). In DC MGs, distributed generation (DG) units are mainly connected to MGs through power electronic interfaces, and the control of parallel DC-DC converters becomes essential (5). The droop control method is the most prevalent control approach employed for DC MGs, which enables cooperative control of power electronic interfaces without the communication link requirements (6). The droop control approach is based on an inner control loop, as a virtual resistance control, which facilitates simultaneous current and power-sharing while offering a plug-and-play feature (7, 8).

The conventional droop method is based on analyzing the output currents and voltages of power electronic interfaces. However, its implementation faces certain challenges. To gain appropriate operation in terms of power sharing, the droop coefficients should be significantly larger than the line resistances. Although the larger values lead to more accurate power sharing, they also create higher voltage drops. Conversely, the smaller values, the better voltage adjustment and the poorer power sharing. As a result, striking a balance between permissible voltage drop and accurate power sharing becomes essential (9). A further issue arises from the fact that line resistances of DG units are mostly different, which adversely impacts both voltage adjustment and power sharing (10).

Various modified droop control methods have been introduced in the existing literature to achieve precise current sharing and voltage regulation in paralleled converters of DC MGs (11-19). It is worth mentioning that modified droop methods are also popular in AC MGs as well (20). An observer-based droop control combined with current feed-forward control was proposed by Li et al. (11) to address the trade-off between the current sharing and the dynamic stability in a DC MG. Both feedback and feed-forward currents were provided by the observer without any additional measurement requirement, resulting in cost reduction. The system's stability, dynamics, and plug-and-play capability were improved as well. Sukhadiaa and Saurabh (12) proposed a droop index for the parallel operation DC converters in low-voltage DC (LVDC) MGs. Their adaptive droop control scheme was based on a function of the normalized current sharing differences of DG units and the power losses, effectively minimizing the circulating current and the current sharing differences. Consequently, the trade-off between voltage regulation and current sharing difference was eliminated. An adaptive droop control algorithm incorporating estimated line resistances was introduced by Ghanbari and Bhattacharya (14), where droop coefficients were set according to these resistances, reducing the line resistance impact on the accurate power sharing. Moreover, a hybrid droop based coordination scheme is recommended by Saeidinia et al. (15) for a class of DC MG that can simultaneously optimize the total generation cost and total transmission power loss. Additionally, a distributed secondary controller was presented to restrain the circulating currents. In the context of interlinking converters connecting AC grids with LVDC subgrids, a selective power droop control approach was proposed by Wang et al. (16)to overcome challenges related to high capacitive coupling impedance and low converter capacity in LVDC grids. The conventional droop strategy was enhanced by Shehata et al. (17) with an online adaptation algorithm that updated the droop coefficients according to the voltage deviation and the current difference. Similarly, a modified droop controller was adopted by Bharath et al. (18), where the droop coefficient was treated as an output voltage-based function, leading to superior voltage adjustment. An adaptive droop control strategy for DC MGs, capable of adapting to the load changes, was established by Shaheed et al. (19). The control approach proposed by Khanabdal et al. (20) exhibited acceptable current sharing performance without requiring any optimization.

In contrast, the previous references did not fully consider the nonlinearity of the system, especially when passive loads, such as resistive ones, are presented in the external characteristics of converters. To address this limitation, the $V_{dc}^2 - P_{dc}$ droop strategy was implemented by Xia et al. (21) for sharing of power and regulating voltage among multiple energy storage units. This decentralized method enhanced system performance and reliability compared to other strategies due to the linear relationship between V_{dc}^2 and P_{dc}, while the relationship between V_{dc} and i_{dc} is nonlinear. Notably, this method is applied here to the primary layer of DG units, which had not been previously considered for the control of DG units further highlighting its outperformance.

The aforementioned droop control methods have not achieved simultaneous precise power sharing and voltage adjustment. Even though the proposed $V_{dc}^2 - P_{dc}$ controller outperforms other primary droop controllers, it still requires a correction term as a secondary controller. Another issue with droop control methods and their derivatives is the lack of a coordinated operation. Therefore, a secondary controller is necessary to address these challenges (22). To achieve this, the centralized, decentralized, or distributed secondary strategies are needed. There are a variety of available approaches, with traditional methods based on centralized control (23) and recent literature focusing on distributed control techniques (24-36). Distributed methods are preferred due to their high reliability and the sparse communication graph requirements.

For instance, a conventional droop-based primary control, along with a feedback-based distributed secondary control, was proposed by Guo et al. (28), for power sharing and voltage restoration in islanded DC MGs. The control apprcoach proposed by Guo et al. (28) required only bus voltages measurment. Aimed to suppress the effects of noise disturbances, a distributed cooperative strategy was introduced by Nabian et al. (29); this secondary controller proposed by Nabian et al. (29) involved two correction terms, one for the voltage and another for the current, and utilized the average voltage of neighboring units. Similar distributed secondary voltage controllers were presented by Biglarahmadi (30) and Nabian et al. (31) for the coordination of DG units in the DC subgrid of hybrid AC/DC MGs, operating under the load changes and the communication link interruption. Guo et al. (33) proposed a distributed attack resilient secondary control method with consensus-based current and average consensus-based voltage regulators, but this method comes with a high computational burden. A two-layer multi-agent based distributed control scheme, comprising physical and cyber voltage restoration layers, is proposed by Fan et al. (34). The control strategy proposed by Fan et al. (34) required only feedback of DC bus voltage and overall system information was not required, however it lacks in-depth discussions on precise power sharing. Another distributed consensus based secondary scheme was suggested Xing et al. (35) for DC MGs in the presence of different loads. This approach utilized a global weighted average voltage in its distributed secondary controller, however, it also involved with a high computational burden. Lastly, Keshavarz et al. (36) considered a new parameter called 'virtual voltage drop' for the secondary control layer, which combined the droop gain and line resistance.

The literature offers various distributed secondary control methods to address challenges in achieving the coordinated operation of DG units and the precise power sharing of them, but each approach comes with its own set of advantages and limitations. The $V_{dc}^2 - P_{dc}$ droop control method addresses the voltage adjustment and the power-sharing (21). Although modifications have been made to the primary droop controllers, achieving accurate power sharing, presice voltage adjustment, and noise disturbance elimination simultaneously remains some challanges. Therefore, the incorporation of a distributed secondary controller becomes essential to accomplish these objectives more effectively. Furthermore, the proposed approach demonstrates enhanced voltage regulation, improved current sharing, and noise nullification compared to data reported in literature (28).

A power sharing improvement term has also been introduced in this study, enabling performance to be improved in scenarios involving communication link interruptions. In summary, the main contributions of this paper are the following:

- The modified distributed secondary control (MDSC) method is designed to effectively cancel noises by approximating the global average voltage through the estimation of local and neighbouring voltages, achieved by nullifying the noise. In comparison with literature (28, 33), this method demonstrates superior noise mitigation capabilities.
- The proposed MDSC method exhibits efficient performance under the load changes, the plug-and-play scenarios, and the communication link malfunctions. This technique demonstrates precise tracking of the reference voltage and achieves nominal voltage with high convergence speed and without overshoot or undershoot, with accurate power sharing as well, as compared to the method described by Keshavarz et al. (36).
- Instead of the conventional droop method, the $V_{dc}^2 - P_{dc}$ approach is employed at the primary control level of DG units. Although this method was previously implemented in voltage adjustment for energy storage units of hybrid MGs (as reported by Xia et al. (21)), it is now applied to the controllers of DG units for the first time. Furthermore, its interaction with the proposed MDSC method is taken into consideration.

The rest of this paper is organized as follows. In section 2, the customized droop method is explained. The proposed secondary control method is discussed in section 3 .The simulation results are presented in section 4 and finally, section 5 concludes this paper.

2. $V_{dc}^2 - P_{dc}$ DROOP METHOD

The general overview of a DC MG with its controllers is demonstrated in Figure 1. As it is shown, each DG unit can be presented with a DC/DC converter connected to photovoltaic (PV) or other generation units. DGs are connected to respective buses with line impedances. The DC MG can be connected to the main grid, even though the understudied DC MG works in islanded mode.

This section presents a $V_{dc}^2 - P_{dc}$ approach for the primary control level. As observed, the conventional

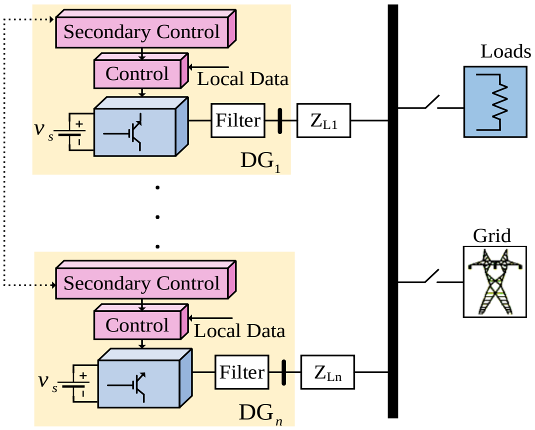


Figure 1. General overview of the DC MG

droop method, which is based on the relationship between voltage and current, needs a precise tradeoff among power sharing accuracy and voltage regulation (28); hence, the droop control approach based on the relationship between the square of voltage and output power designed by Xia et al. (21) and implemented for energy storage units is considered for the primary control level of DG units in this paper. The typical diagram of a DG which is connected by a DC/DC converter to the DC common bus is shown in Figure 2. According to this figure, the relationship between the output current and voltage ($v_o - i_o$) can be expressed as folows (21):

$$\begin{cases} c_f \dfrac{dv_c}{dt} = i_{in} - i_o , \\ v_c = v_o , \\ i_o = v_o R_{load} + \dfrac{P_{DC}}{v_o} \end{cases} \tag{1}$$

In the above equation, c_f, R_{load}, P_{DC}, i_{in}, and v_c represent filter capacitance, load resistance, DC power, input current, and capacitance voltage, respectively. As it is obvious, the relationship of $v_o - i_o$ is nonlinear which means if the traditional droop method is employed, the efficiency of regulations will be reduced.

Despite all that, the forthcoming equations can be written if the balance of power in Figure 2 is taken into account (21).

$$\left.\begin{aligned} c_f \frac{dv_o^2}{dt} &= P_{in} - P_o \\ P_o &= \frac{v_o^2}{R_{load}} + P_{DC} \end{aligned}\right\} \Rightarrow c_f \frac{dv_o^2}{dt} = P_{in} - \frac{v_o^2}{R_{load}} - P_{DC} \tag{2}$$

where P_{in} and P_o are the input and the output powers. It is obvious that there is a linear relationship between v_o^2 and P_o. As a result, a droop method based on $P - v_o^2$ can be derived in order to improve the accuracy of the control system. This method is composed of two inner and outer loops, the former controlling the v_o^2 and the later considering the power sharing.

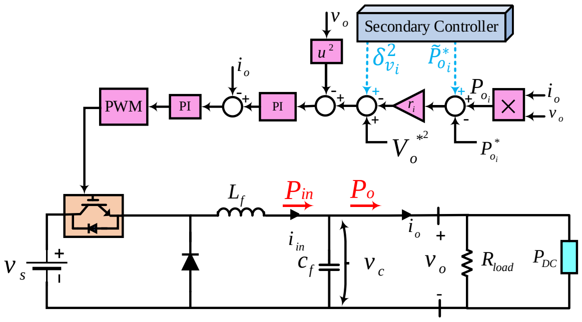


Figure 2. The typical diagram of a DC DG with primary controller and its simplified circuit diagram with connected loads

If the switch and filter losses are neglected, the average model of the converter can be written as follows (21):

$$\begin{cases} \dfrac{1}{2} c_f \dfrac{dv_o^2}{dt} = v_s . i_{in} - P_o \\ L_f \dfrac{di_{in}}{dt} = v_s - v_o \end{cases} \tag{3}$$

where L_f and v_s represent the filter inductance and the voltage of primary resource, respectively. With regard to this equation, the inner loop can be illustrated as Figure 2. A proportional-integral (PI) controller for the voltage controller is employed in order to eliminate static errors, while a proportional controller is utilized for the current controller in order to intensify damping. Next, the customized $P_o - v_o^2$ o droop, as outer loop, can be analyzed and formulized as follows (21).

$$v_{o_i}^{ref^2} = V_o^{*2} - r_i \left(P_{o_i} - P_{o_i}^* \right) \tag{4}$$

where $v_{o_i}^{ref^2}$, V_o^{*2}, r_i, P_{o_i}, and $P_{o_i}^*$ represent reference output voltage value, rated output voltage value, droop coefficient, measured output power, and rated output power, respectively; subscript of i denotes the i^{th} converter. Besides, the following equation should be satisfied if it is required to ensure appropriate current sharing while load changes (21).

$$\frac{r_i}{r_j} = \frac{P_{o_j}^*}{P_{o_i}^*} \tag{5}$$

3. DISTRIBUTED SECONDARY CONTROL FOR VOLTAGE RESTORATION

In this part, the proposed MDSC method, which is depicted in Figure 3(a), is explained. As a result, the following term is regarded as the secondary correction voltage, and is written in the following manner:

$$\delta_{v_i} = \left(k_p + \frac{k_i}{s}\right)e_i \tag{6}$$

where k_p and k_i represent the PI parameters and e_i is total error, which can be expressed as follows:

$$e_i = g_i\ \alpha_i\ e^v + \beta_i e^{\delta_{v_i}} \tag{7}$$

In the above equation, α_i, β_i, and g_i are control error and restoration error gains and pinning gain ($g_i \neq 0$ for the reference agent), respectively; $e^{\delta_{v_i}}$ denotes the control error and e^v represents the error of restoring voltage. These recent terms are defined as follows:

$$\begin{cases} e^v = v_o^* - v_o \\ e^{\delta_{v_i}} = \sum_{j\in N} a_{ij}\left(\delta_{v_j} - \delta_{v_i}\right) \end{cases} \tag{8}$$

where a_{ij} is the edge weight among i th and j th DG. As one of the aims of this controller is voltage restoration, then:

$$\begin{cases} \lim_{t\to\infty} e^v(t) = 0 \\ \lim_{t\to\infty} e^{\delta_{v_i}}(t) = 0 \end{cases} \tag{9}$$

The performance of the proposed voltage controller can be influenced if disturbances exist; hence, a term should be considered for nullification of noises. This term is defined as follows:

$$u_i = v_i - \delta_{v_i} \tag{10}$$

A white noise source, which is illustrated by d_i in Figure 3 is exerted. To do this and ensure accurate disturbance tracking, u_i is sent to an integrator. Afterward, the average value of this term can be rewritten as follows:

$$\dot{\bar{u}}_i(t) = u_i(t) + \int_0^t \sum_{j\in N} ai_j\left(\bar{u}_j(\tau) - \bar{u}_i(\tau)\right)d\tau \tag{11}$$

The following protocol is attained by differentiation of Equation 11.

$$\dot{\bar{u}}_i(t) = u_i(t) + \sum_{j\in N} ai_j\left(\bar{u}_j - \bar{u}_i\right) \tag{12}$$

Eventually, $\tilde{d}_i$, as the noise deactivation term, is defined as follows:

$$\tilde{d}_i = \frac{k_i}{s}\bar{u}_i \tag{13}$$

In the end, the new forthcoming term is assumed as the reference voltage value for the primary controller.

$$\delta_{v_i}^{'2} = \left(k_p + \frac{k_i}{s}\right)e_i + d_i - \tilde{d}_i \tag{14}$$

The other aim of the secondary controller is to make the output power track its reference value meticulously. As a result, a PI controller is adopted because of its capability to omit the steady state error as follows:

$$\tilde{P}_{o_i}^* = \left(k_{P_p} + \frac{k_{P_i}}{s}\right)\left(P_{o_i}^* - P_{o_i}\right) \tag{15}$$

where k_{P_p} and k_{P_i} represent the PI coefficients and $\tilde{P}_{o_i}^*$ is produced reference power, which is named as virtual power control reference. It will be utilized as the reference value of the primary controller. For achieving better and more accurate current sharing, the following term is also added by considering the output powers and the droop coefficients of other DGs.

$$\delta_p = \sum_{j\in N} a_{ij}\left(r_j P_{o_j} - r_i P_{o_i}\right) \tag{16}$$

where δ_p denotes the power correction term. The final virtual power reference ($\tilde{P}_{o_i}^*$) which is added to $P_{o_i}^*$ in the primary controller can be rewritten as follows:

$$\tilde{P}_{o_i}^* = \left(k_{P_p} + \frac{k_{P_i}}{s}\right)\left(P_{o_i}^* - P_{o_i}\right) + \delta_p \tag{17}$$

The overall structure of this controller is illustrated in Figure 3(b).

4. SIMULATION RESULTS

For validating the proposed CDM and MDSC methods in DC MGs, a typical DC MG which is depicted in Figure 4 is simulated in MATLAB/SIMULINK software. This MG is composed of 4 DG units and 800W loads; the

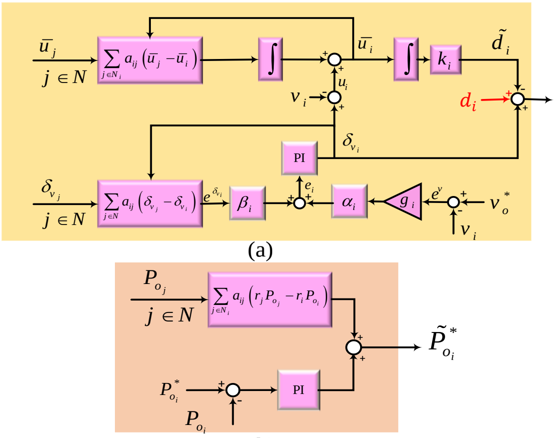


Figure 3. Control block diagram of the MDSC method (a) the voltage correction term (b) the power correction term $\delta_{v_i}^{'2}$ is added to the primary controller (the blue one in Figure 2)

nominal produced power of DGs is considered as 200W. The parameters of DGs, lines, and other data can be found in Table 1. To confirm the efficiency of the proposed method, different case studies are considered and results are also compared with other methods, including the conventional primary droop method and proposed distributed secondary control methods (28, 35) for better evaluation. The case studies are as follows:

- Case I: The DC MG works with either conventional droop or $V_{dc}^2 - P_{dc}$ droop method;
- Case II: At *t = 1.5s*, the voltage restoration part of the MDSC method and the methods of Guo et al. (28) and Keshavarz et al. (36) are activated for each respective simulation;
- Case III: The power correction term of the MDSC method is subsequently added at *t = 2.5s*, so its efficiency is more sensible;
- Case IV: There is a load increase at *t = 3.5s* (its power equals 200W);
- Case V: DG$_4$ is plugged out at *t = 4.5s*, and reconnected at *t = 6s*;
- Case VI: At *t = 7.5s*, the communication link between DG$_2$ and DG$_3$ is missed and it is restored at *t = 9s*;
- Case VII: Finally, a white noise is considered at *t = 10.5s*.

It is worth mentioning that a graph should be assumed to demonstrate the communication links among DGs. The following equation shows the corresponding adjacency matrix. According to this matrix, there is a unilateral connection which reduces the volume of data.

$$A = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \tag{18}$$

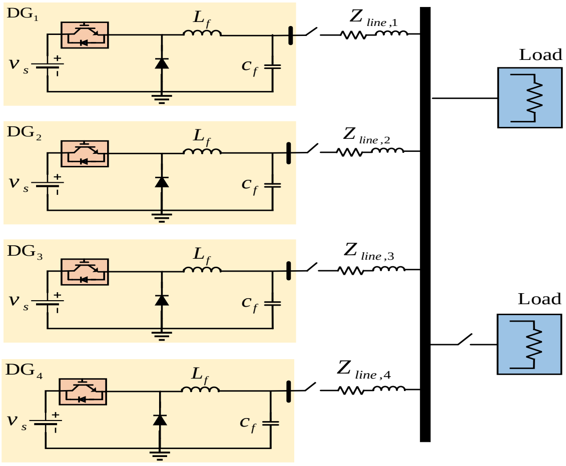


Figure 4. Structure of the studied DC MG

TABLE 1. Parameters of DC MG

Symbol	Quantity	Value
V_o	Nominal DC voltage	48 V
k_{p_v}	Proportional term of voltage controller	1
k_{i_v}	Integral term of voltage controller	250
k_{p_i}	Proportional term of current controller	5
k_{i_i}	Integral term of voltage controller	100
$r_{1\&2}$	DG1&DG2 Droop coefficients	$5\ V^2/W$
$r_{3\&4}$	DG3&DG4 Droop coefficients	$3\ V^2/W$
	Secondary parameters	
α_i	control error gain	1
β_i	restoration error gain	1
k_i	Integral term of secondary voltage controller	1
$k_{P\&I}$	Proportional term of secondary power controller	1 & 40

The obtained results are presented in Figures 5-10 as follows.

The results of Cases I-IV are illustrated in Figures 5(a)-(f) and 6(a)-(f). As it is obvious, when the primary controllers are merely activated, the proposed $V_{dc}^2 - P_{dc}$ droop method shows superior performance (Figures 5 and 6 (b)) and the voltage drop is lower in comparison with results obtained by the conventional droop controller (Figure 5 and 6(a)). As long as the secondary method of (28, 36) and the proposed MDSC method are employed in respective simulations, although both strategies restore the voltage to its nominal value, no overshoot can be observed for implementing the MDSC method and it also converges faster. In addition, when the correction power term of the MDSC method is implemented (*t = 2.5s*), power and current sharing are improved subsequently (see Figures 5 and 6(d) & (f)). Moreover, while load changes, it is manifest that the voltage drop of proposed methods is ignorable, while there is an observable drop for voltage obtained by the method of Keshavarz et al. (36). For a more precise comparison, Table 2 provides the voltage error percentages for each case. The time required to approach nominal values is compared in Table 3. It should be mentioned that both simulations are performed in similar situations and both by 8 GB RAM, Core i7-2670QM CPU system.

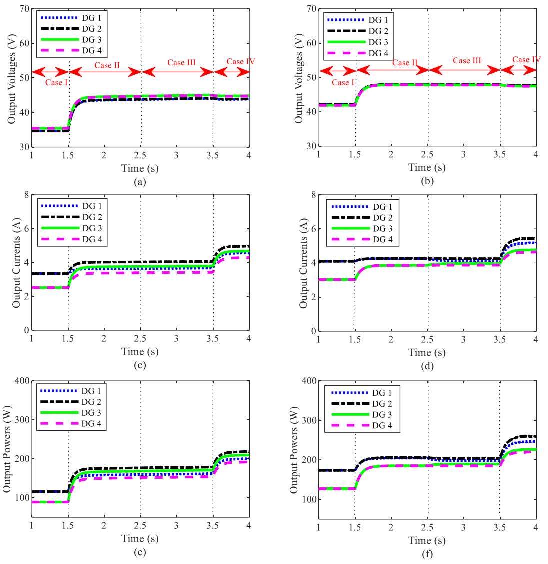


Figure 5. Results obtained in Cases I-IV. (a) Output voltage obtained by method of (28); (b) Output voltage obtained by proposed $V_{dc}^2 - P_{dc}$ droop and MDSC methods; (c) Output current obtained by method of (28); (d) Output current obtained by proposed $V_{dc}^2 - P_{dc}$ droop and MDSC methods; (e) Output power obtained by method of (28); (f) Output power obtained by proposed $V_{dc}^2 - P_{dc}$ droop and MDSC methods

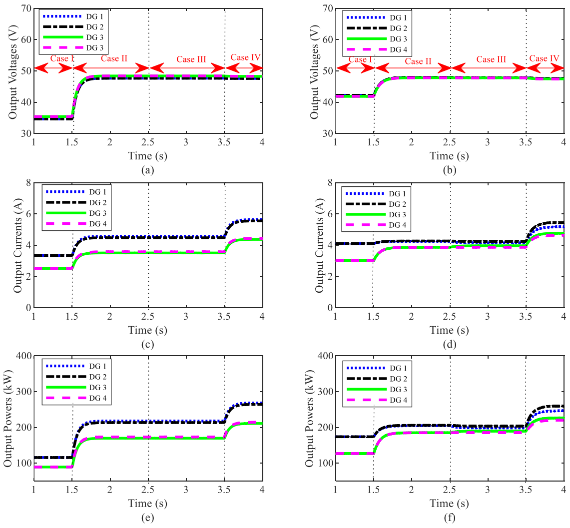


Figure 6. Results obtained in Cases I-IV. (a) Output voltage obtained by method of (36); (b) Output voltage obtained by proposed $V_{dc}^2 - P_{dc}$ droop and MDSC methods; (c) Output current obtained by method of (36); (d) Output current obtained by proposed $V_{dc}^2 - P_{dc}$ droop and MDSC methods; (e) Output power obtained by method of (36); (f) Output power obtained by proposed $V_{dc}^2 - P_{dc}$ droop and MDSC methods

TABLE 2. Comparison of errors in different cases

Index	$e_{V_i}(\%)=\frac{V_{ref}-V_i}{V_{ref}}\times 100$							
	Case	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Conventional droop and method of (28)	DG$_1$	27.91	7.1	-	6.69	25.68	6.9	6.04
	DG$_2$	27.91	7.05	-	6.65	23.23	6.76	5.83
	DG$_3$	26.25	8.93	-	8.52	19.37	4.79	3.75
	DG$_4$	26.25	8.93	-	8.56	-	5	3.95
Conventional droop and method of (36)	DG$_1$	27.91	0.416	-	1.341	2.083	1.666	-6.73
	DG$_2$	27.91	0.416	-	1.341	2.083	1.145	-6.73
	DG$_3$	26.25	0.833	-	1.537	0.83	2.083	-7.41
	DG$_4$	26.25	0.833	-	1.537	-	1.145	-7.41
$V_{dc}^2 - P_{dc}$ droop and MDSC methods	DG$_1$	10.83	0.145	-	0.719	1.666	0.729	0.729
	DG$_2$	10.83	0.145	-	0.719	1.666	0.583	0.729
	DG$_3$	12.72	0.354	-	0.921	1.395	1.145	0.937
	DG$_4$	12.72	0.354	-	0.921	-	0.583	0.937
Index	$e_{P_i}(\%)=\left(\frac{P_i}{\sum P_i}-\frac{P_{rated}}{\sum P_{rated}}\right)\times 100$							
	Case	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Conventional droop and method of (28)	DG$_1$	5.048	-0.65	-	-0.31	-18.04	-1.02	-1.28
	DG$_2$	5.048	1.8	-	0.8	9.78	1.43	0.65
	DG$_3$	-5.048	0.88	-	0.93	33.26	0.73	1.22
	DG$_4$	-5.048	-2.03	-	-1.42	-	-1.14	-0.59
Conventional droop and method of (36)	DG$_1$	5.048	2.272	-	3.010	8.96	3.125	2.574
	DG$_2$	5.048	2.272	-	3.010	8.96	3.125	6.412
	DG$_3$	-5.048	-2.272	-	-3.010	10.507	-3.125	-3.5
	DG$_4$	-5.048	-2.272	-	-3.010	-	-3.125	-3.5
$V_{dc}^2 - P_{dc}$ droop and MDSC methods	DG$_1$	3.903	1.282	0.410	0.941	3.295	-0.376	0.492
	DG$_2$	3.903	1.282	1.089	1.696	3.295	-0.376	0.492
	DG$_3$	-3.903	-1.282	-0.641	-1.15	3.077	-1.188	-1.188
	DG$_4$	-3.903	-1.282	-1.282	-1.987	-	-1.188	-1.542

TABLE 3. Comparison of time to reach nominal values (s)

	Ref.	**[28]**	**[35]**	**This paper**
After activation of Sec. Controller	V	0.23	0.15	0.15
	I	0.2	0.1	0.1
	P	0.22	0.2	0.2
After applying noise	V	1.45	0.9	0.01
	I	1.425	1	0.01
	P	1.985	1	0.02

Results for the efficiency evaluation of proposed methods in plug-and-play condition (Case V) are depicted in Figures 7 (a)-(f) and 8(a)-(f). It can be deemed that when DG_4 is disconnected, the method of Guo et al. (28) is unable to restore voltages and it takes some seconds to restore the voltage by using the method of Keshavarz et al. (36) (Figures 7 and 8(a)); however, the proposed methods properly remained stable (Figures 7 and 8(b)). Besides, when DG_4 is reconnected, it is completely clear that there is a significant overshoot for results obtained by the method of Keshavarz et al. (36), and even after 1s, it is not able to restore the stable condition in the studied MG (Figure 8(c)); by the method of Guo et al. (28), it takes more than *5s* to become stable (Figures 7(c)), though the results achieved by proposed strategies appropriately restore the voltage and other parameters without any overshoot (Figures 7 and 8(d)). The respective errors are given in Table 2. Eventually, Figures 9 (a)-(f) and 10 (a)-(f) show the results for Cases VI and VII. It can be observed that when the communication link between DG_2 and DG_3 is lost, there is a conspicuous drop in the voltage and power of DG_2 by methods of Guo et al. (28) and Keshavarz et al. (36). Since white noise is inserted, despite methods of (28, 36), the proposed strategies work efficiently in the existence of noise and remain stable. To compare the results properly, the respective error percentages are presented in Table 2. It is obvious from the table that when the presented control methods are considered, the

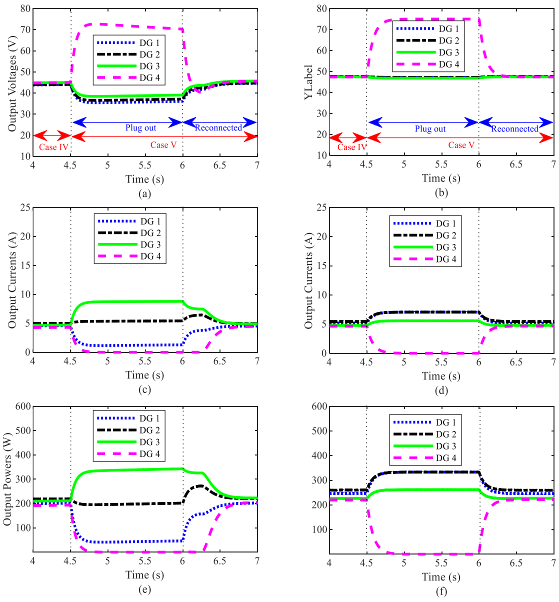


Figure 7. Results obtained in Cases V. (a) Output voltage obtained by method of (28); (b) Output voltage obtained by proposed $V_{dc}^2 - P_{dc}$ droop and MDSC methods; (c) Output current obtained by method of (28); (d) Output current obtained by proposed $V_{dc}^2 - P_{dc}$ droop and MDSC methods; (e) Output power obtained by method of (28); (f) Output power obtained by proposed $V_{dc}^2 - P_{dc}$ droop and MDSC methods

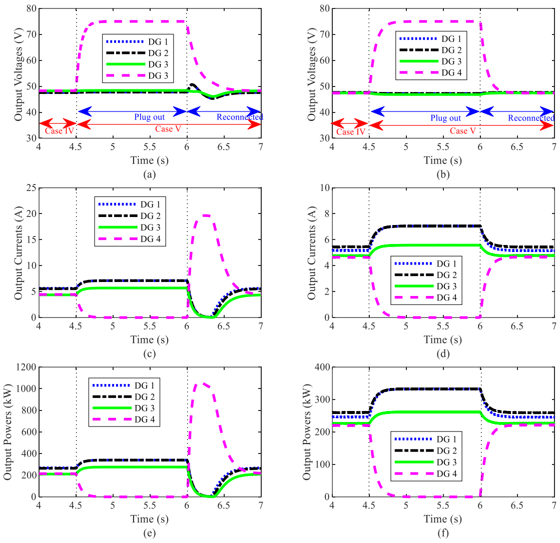


Figure 8. Results obtained in Cases V. (a) Output voltage obtained by method of (36); (b) Output voltage obtained by proposed $V_{dc}^2 - P_{dc}$ droop and MDSC methods; (c) Output current obtained by method of (36); (d) Output current obtained by proposed $V_{dc}^2 - P_{dc}$ droop and MDSC methods; (e) Output power obtained by method of (36); (f) Output power obtained by proposed $V_{dc}^2 - P_{dc}$ droop and MDSC methods

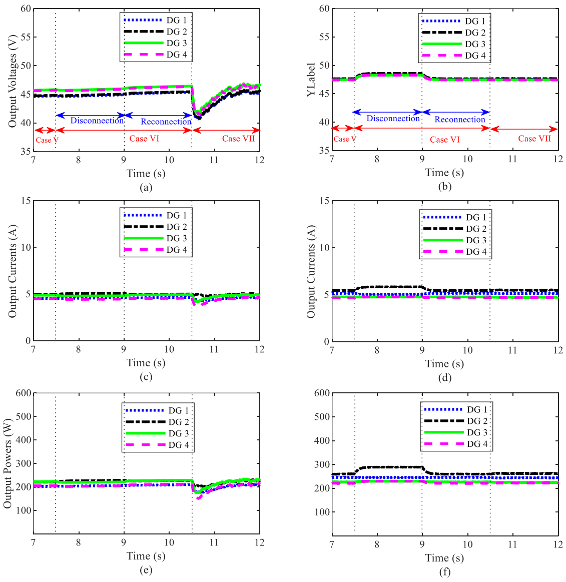


Figure 9. Results obtained in Cases VI-VII. (a) Output voltage obtained by method of (28); (b) Output voltage obtained by proposed $V_{dc}^2 - P_{dc}$ droop and MDSC methods; (c) Output current obtained by method of (28); (d) Output current obtained by proposed $V_{dc}^2 - P_{dc}$ droop and MDSC methods; (e) Output power obtained by method of (28); (f) Output power obtained by proposed $V_{dc}^2 - P_{dc}$ droop and MDSC methods

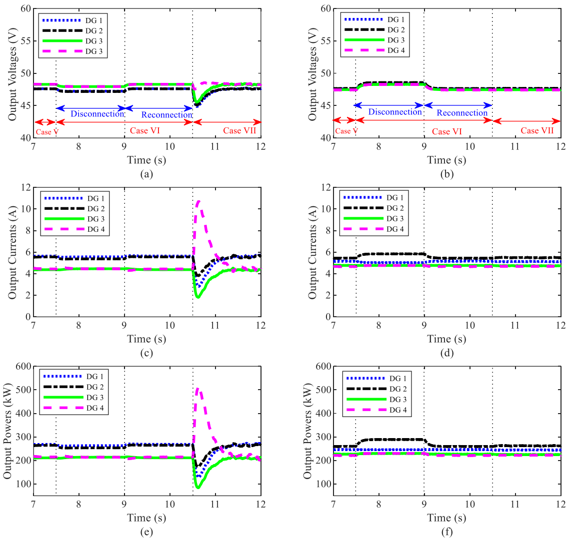


Figure 10. Results obtained in Cases VI-VII. (a) Output voltage obtained by method of (36); (b) Output voltage obtained by proposed $V_{dc}^2 - P_{dc}$ droop and MDSC methods; (c) Output current obtained by method of (36); (d) Output current obtained by proposed $V_{dc}^2 - P_{dc}$ droop and MDSC methods; (e) Output power obtained by method of (36); (f) Output power obtained by proposed $V_{dc}^2 - P_{dc}$ droop and MDSC methods

errors are in the acceptable range, while by implementing other methods (28, 36), there are considerable errors. The conventional droop-based primary controller along with the method of Guo et al. (28) is unable to share power and regulate voltage, especially when a disturbance such as plugging out of DG occurs.

Additionally, when the secondary proposed controller is activated, the time needed to reach the desired valuses are to some extent similar based on Table 3. However, if the noise is applied, the method of this paper can more effectively perform and has comparable better results compared to the methods of (28, 36). As observed, the strategy of Guo et al. (28) requires more than *1.5 s* to overcome this challenge; that of Keshavarz et al. (36) is almost *1 s*, while the proposed scheme approaches desirable values less than *0.05 s*.

5. CONCLUSION

This paper introduces a two-layer control structure aimed at effectively managing voltage control, accurately sharing current, and precisely performing across various operational scenarios as well as noise existence. To achieve this, a customized droop strategy was implemented in the first layer, resulting in reduced voltage drop and improved power sharing in comparison with conventional droop strategies. Furthermore, a distributed secondary controller incorporating voltage and power correction terms was integrated into the primary level to regulate the voltage at its nominal value and enhance precision in power allocation. Comparative analysis was conducted, revealing that the proposed $V_{dc}^2 - P_{dc}$ droop and MDSC methods outperformed the examined case studies in terms of less overshoot and faster convergence. Additionally, the effectiveness of the proposed methods was demonstrated when subjected to white noise injection.

6. REFERENCES

1. Sagar G, Debela T. Implementation of optimal load balancing strategy for hybrid energy management system in dc/ac microgrid with pv and battery storage. International Journal of Engineering, Transactions A: Basics. 2019;32(10):1437-45. 10.5829/IJE.2019.32.10A.13
2. Khanalizadeh Eini M, Mirhosseini Moghadam M, Tavakoli A, Alizadeh B. Stability Analysis of AC/DC Microgrids in Island Mode. International Journal of Engineering, Transactions A: Basics. 2021;34(7):1750-65. 10.5829/IJE.2021.34.07A.20
3. Norozpour Niazi A, Vasegh N, Motie Birjandi A. An Improved Hierarchical Control Structure for Robust Microgrid Operation and Seamless Mode Transfer under Linear and Nonlinear Loads conditions. International Journal of Engineering, Transactions C: Aspects. 2021;34(9):2167-79. 10.5829/IJE.2021.34.09C.14
4. Chen Y-K, Wu Y-C, Song C-C, Chen Y-S. Design and implementation of energy management system with fuzzy control for DC microgrid systems. IEEE Transactions on power electronics. 2012;28(4):1563-70. 10.1109/TPEL.2012.2210446
5. Abdolahi M, Adabi J, Mousavi SYM. An Adaptive Extended Kalman Filter with Passivity-Based Control for DC-DC Converter in DC Microgrids Supplying Constant Power Loads. IEEE Transactions on Industrial Electronics. 2023. 10.1109/TIE.2023.3283686
6. Nejabatkhah F, Li YW. Overview of power management strategies of hybrid AC/DC microgrid. IEEE Transactions on power electronics. 2014;30(12):7072-89. 10.1109/TPEL.2014.2384999
7. Dragičević T, Lu X, Vasquez JC, Guerrero JM. DC microgrids—Part II: A review of power architectures, applications, and standardization issues. IEEE transactions on power electronics. 2015;31(5):3528-49. 10.1109/TPEL.2015.2464277
8. Khorsandi A, Ashourloo M, Mokhtari H. A decentralized control method for a low-voltage DC microgrid. IEEE Transactions on Energy Conversion. 2014;29(4):793-801. 10.1109/TEC.2014.2329236
9. Elrayyah A, Cingoz F, Sozer Y. Smart loads management using droop-based control in integrated microgrid systems. IEEE Journal of Emerging and Selected Topics in Power Electronics. 2017;5(3):1142-53. 10.1109/JESTPE.2017.2666786
10. Kumar R, Pathak MK. Distributed droop control of dc microgrid for improved voltage regulation and current sharing. IET Renewable Power Generation. 2020;14(13):2499-506. 10.1049/iet-rpg.2019.0983
11. Li X, Guo L, Zhang S, Wang C, Li YW, Chen A, et al. Observer-based DC voltage droop and current feed-forward control of a DC microgrid. IEEE Transactions on Smart Grid. 2017;9(5):5207-16. 10.1109/TSG.2017.2684178

12. Sukhadiaa R, Pandya S, editors. Modified Droop Control Strategy for Load Sharing and Circulating Current Minimization in Low-Voltage Standalone DC Microgrid. Advances in Control Systems and its Infrastructure: Proceedings of ICPCCI 2019; 2020: Springer. 10.1007/978-981-15-0226-2_5
13. Zhao W, Duan M, Dong Y, Wang G, Pan Y, Tang M, et al., editors. An Adaptive Dynamic Droop Control Strategy for Load Sharing and Voltage Regulation in Low-Voltage DC Microgrid. 2021 IEEE Sustainable Power and Energy Conference (iSPEC); 2021: IEEE. 10.1109/iSPEC53008.2021.9735666
14. Ghanbari N, Bhattacharya S. Disturbance rejection analysis of a droop-controlled dc microgrid through a novel mathematical modeling. IEEE Journal of Emerging and Selected Topics in Power Electronics. 2021;10(2):1507-18. 10.1109/JESTPE.2021.3088777
15. Saeidinia Y, Arabshahi M, Mousazadeh Mousavi SY, Biglari M. Autonomous control of DC microgrid based on a hybrid droop control scheme for total generation cost and transmission power loss reduction. Electrical Engineering. 2023;105(1):267-83. 10.1007/s00202-022-01662-w
16. Wang L, Wong M-C, Zhou X, He Z, Xu Q, Zhou L. A selective power droop control for hybrid interlinking converter in AC/LVDC microgrid. IEEE Transactions on Industrial Electronics. 2020;68(10):9046-57. 10.1109/TIE.2020.3020017
17. Shehata E, Thomas J, Mostafa R, Ghalib M, editors. An improved droop control for a low voltage DC microgrid operation. 2018 Twentieth International Middle East Power Systems Conference (MEPCON); 2018: IEEE. 10.1109/MEPCON.2018.8635258
18. Bharath K, Dayal A, Kanakasabapathy P, editors. A simulation study on modified droop control for improved voltage regulation in DC microgrid. 2017 International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT); 2017: IEEE. 10.1109/ICICICT1.2017.8342581
19. Shaheed MNB, Sozer Y, Chowdhury S, De Abreu-Garcia JA, editors. A novel decentralized adaptive droop control technique for DC microgrids based on integrated load condition processing. 2020 IEEE Energy Conversion Congress and Exposition (ECCE); 2020: IEEE. 10.1109/ECCE44975.2020.9235785
20. Khanabdal S, Banejad M, Blaabjerg F, Hosseinzadeh N. A novel control strategy of an islanded microgrid based on virtual flux droop control and direct flux fuzzy control. International Journal of Engineering, Transactions B: Applications,. 2021;34(5): 1274-83. 10.5829/IJE.2021.34.05B.21
21. Xia Y, Wei W, Yu M, Wang X, Peng Y. Power management for a hybrid AC/DC microgrid with multiple subgrids. IEEE Transactions on power electronics. 2017;33(4):3520-33. 10.1109/TPEL.2017.2705133
22. Dragicevic T, Vasquez JC, Guerrero JM, Skrlec D. Advanced LVDC electrical power architectures and microgrids: A step toward a new generation of power distribution networks. IEEE Electrification Magazine. 2014;2(1):54-65. 10.1109/MELE.2013.229703
23. Guerrero JM, Vasquez JC, Matas J, De Vicuña LG, Castilla M. Hierarchical control of droop-controlled AC and DC microgrids—A general approach toward standardization. IEEE Transactions on industrial electronics. 2010;58(1):158-72. 10.1109/TIE.2010.2066534
24. Nasirian V, Moayedi S, Davoudi A, Lewis FL. Distributed cooperative control of DC microgrids. IEEE Transactions on Power Electronics. 2014;30(4):2288-303. 10.1109/TPEL.2014.2324579
25. Anand S, Fernandes BG, Guerrero J. Distributed control to ensure proportional load sharing and improve voltage regulation in low-voltage DC microgrids. IEEE transactions on power electronics. 2012;28(4):1900-13. 10.1109/TPEL.2012.2215055
26. Shafiee Q, Dragicevic T, Andrade F, Vasquez JC, Guerrero JM, editors. Distributed consensus-based control of multiple DC-microgrids clusters. IECON 2014-40th Annual Conference of the IEEE Industrial Electronics Society; 2014: IEEE. 10.1109/IECON.2014.7048785
27. Meng L, Dragicevic T, Guerrero JM, Vasquez JC, editors. Dynamic consensus algorithm based distributed global efficiency optimization of a droop controlled DC microgrid. 2014 IEEE international energy conference (ENERGYCON); 2014: IEEE. 10.1109/ENERGYCON.2014.6850587
28. Guo F, Xu Q, Wen C, Wang L, Wang P. Distributed secondary control for power allocation and voltage restoration in islanded DC microgrids. IEEE Transactions on Sustainable Energy. 2018;9(4):1857-69. 10.1109/TSTE.2018.2816944
29. Dehaghani MN, Taher SA, Arani ZD, editors. Distributed Secondary Voltage and Current Control Scheme with Noise Nullification Ability for DC Microgrids. 2020 10th Smart Grid Conference (SGC); 2020: IEEE. 10.1109/SGC52076.2020.9335768
30. Biglarahmadi M, Ketabi A, Baghaee HR, Guerrero JM. Integrated nonlinear hierarchical control and management of hybrid AC/DC microgrids. IEEE Systems Journal. 2021;16(1):902-13. 10.1109/JSYST.2021.3050334
31. Nabian Dehaghani M, Taher SA, Dehghani Arani Z. An efficient power sharing approach in islanded hybrid AC/DC microgrid based on cooperative secondary control. International Transactions on Electrical Energy Systems. 2021;31(6):e12897. 10.1002/2050-7038.12897
32. Zuo S, Altun T, Lewis FL, Davoudi A. Distributed resilient secondary control of DC microgrids against unbounded attacks. IEEE Transactions on Smart Grid. 2020;11(5):3850-9. 10.1109/TSG.2020.2992118
33. Guo F, Wang L, Wen C, Zhang D, Xu Q. Distributed voltage restoration and current sharing control in islanded DC microgrid systems without continuous communication. IEEE Transactions on Industrial Electronics. 2019;67(4):3043-53. 10.1109/TIE.2019.2907507
34. Fan B, Peng J, Yang Q, Liu W. Distributed control of DC microgrids with improved ZIP load adaptability. IEEE Transactions on Systems, Man, and Cybernetics: Systems. 2021;52(7):4623-33. 10.1109/TSMC.2021.3101813
35. Xing L, Mishra Y, Guo F, Lin P, Yang Y, Ledwich G, et al. Distributed secondary control for current sharing and voltage restoration in DC microgrid. IEEE Transactions on Smart Grid. 2019;11(3):2487-97. 10.1109/TSG.2019.2956515
36. Keshavarz M, Doroudi A, Kazemi M, Mahdian Dehkordi N. A new consensus-based distributed adaptive control for islanded microgrids. International Journal of Engineering, Transactions A: Basics. 2021;34(7):1725-35. 10.5829/IJE.2021.34.07A.17

Persian Abstract

چکیده

در سال‌های اخیر، ریزشبکه‌های DC به سبب افزایش گسترده منابع انرژی پراکنده و بارهاDC ، به همراه پیشرفت در فناوری‌های الکترونیک قدرت، توجه قابل توجهی را به خود جلب کرده‌اند. به منظور به‌کارگیری مؤثر ریزشبکه‌ها، استفاده از یک استراتژی کنترلی مناسب که ولتاژ ریزشبکه را در سطح مطلوب حفظ کند و تقسیم توان برابر بین واحدهای تولید پراکنده را تضمین کند، بسیار مهم است. جهت دستیابی به این اهداف، یک ساختار کنترلی دو لایه در این مقاله پیشنهاد شده است. در لایه اولیه، یک طرح کنترل شیب–افتی اصلاحی معرفی شده است که افت ولتاژ کمتری را در مقایسه با استراتژی‌های شیب–افتی سنتی به دست می‌آورد. همچنین، یک کنترل‌کننده توزیع‌شده بهبود دهنده ولتاژ در لایه ثانویه پیشنهاد شده است که دارای یک قسمت کمکی جهت بهبود دقت تقسیم توان نیز می‌باشد. استراتژی‌های کنترل پیشنهادی قابلیت اجرای ساده‌ای دارند و به شبکه ارتباطی و مخابراتی کم حجمی نیاز دارند. برای ارزیابی کارایی روش کنترل پیشنهادی، مطالعات موردی متعددی مانند تغییرات بار و چالش‌های قطع و وصل شبکه ارتباطی و اثرات نویز انجام شده‌است. علاوه بر این، عملکرد روش پیشنهادی در برابر طرح کنترل افتی اولیه سنتی و روش‌های کنترل ثانویه توزیع‌شده مشارکتی مقایسه شده است. مطالعات شبیه‌سازی انجام‌شده در نرم‌افزار MATLAB/SIMULINK نشان می‌دهد که روش کنترل پیشنهادی بهتر از روش‌های دیگر عمل می‌کنند و اثربخشی آن را در حفظ تنظیم ولتاژ و اهداف اشتراک توان تأیید می‌کند.

IJE TRANSACTIONS B: Applications Vol. 37 No. 02, (February 2024) 352-364

International Journal of Engineering

Journal Homepage: www.ije.ir

A Comparative Study of Hybrid Analytical and Laplace Transform Approaches for Solving Partial Differential Equations in Python

P. Jalili[a], M. Mahboob[b], A. Shateri[a], B. Jalili[a], D. Domiri Ganji*[b]

[a] *Department of Mechanical Engineering, North Tehran Branch, Islamic Azad University, Tehran, Iran*
[b] *Department of Mechanical Engineering, Babol Noshirvani University of Technology, Babol, Iran*

PAPER INFO

Paper history:
Received 25 August 2023
Received in revised form 25 September 2023
Accepted 26 September 2023

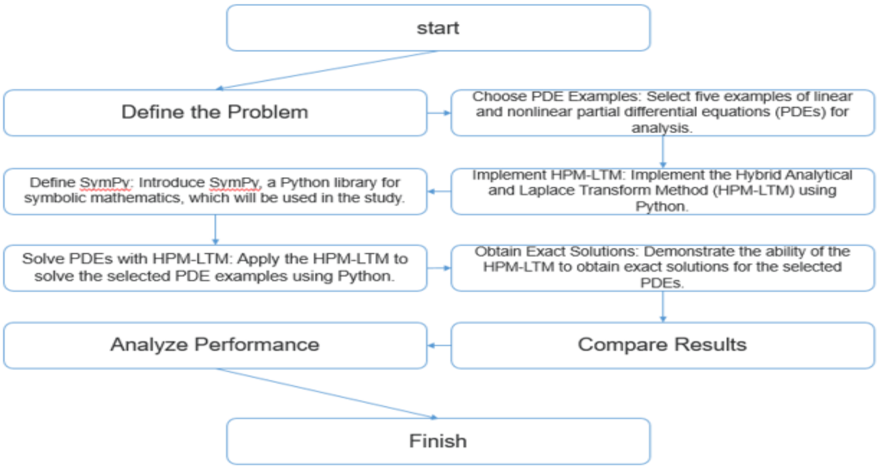
Keywords:
Partial Differential Equations
Homotopy Perturbation Method
Laplace Transform Method
Python Programming

ABSTRACT

This research presents a rigorous and innovative approach, the Homotopy Perturbation Method-Laplace Transform Method (HPM-LTM), implemented in Python, for the efficient solution of linear and nonlinear partial differential equations (PDEs). By combining the Homotopy Perturbation technique (HPM) with the Laplace Transform Method (LTM), our method successfully addresses the significant challenges posed by equations with nonlinear components. Through the utilization of He's polynomials, the HPM-LTM approach effectively handles nonlinear terms, resulting in accurate and reliable solutions. To demonstrate the efficacy of our method, we extensively apply it to five representative PDE scenarios, including heat and wave equations. Our comprehensive results substantiate the remarkable accuracy and reliability of the HPM-LTM approach, highlighting its superiority in comparison to conventional approaches that require restrictive assumptions or discretization, which can introduce round-off errors. Furthermore, our method overcomes the limitations imposed by numerical errors inherent in traditional HPM techniques. The robustness, effectiveness, and adaptability of our proposed approach are further validated by its successful application to a wide range of PDE problems across various fields. This research presents a significant contribution to the development of a powerful computational tool for resolving diverse PDE problems, with particular relevance to the engineering discipline.

doi: *10.5829/ije.2024.37.02b.11*

Graphical Abstract



---

*Corresponding Author Email: *mirgang@nit.ac.ir* (D. Domiri Ganji)

Please cite this article as: Jalili P, Mahboob M, Shateri A, Jalili B, Domiri Ganji D. A Comparative Study of Hybrid Analytical and Laplace Transform Approaches for Solving Partial Differential Equations in Python, International Journal of Engineering, Transactions B: Applications. 2024;37(02):352-64.

1. INTRODUCTION

Partial differential equations (PDEs) are fundamental for modeling various physical phenomena, such as fluid dynamics, heat transfer, and electromagnetic fields. Analytic solutions for PDEs are usually not available, leading to the need for numerical methods to obtain approximate solutions. The Homotopy Perturbation Method (HPM) has recently become an efficient and reliable method for solving linear and nonlinear PDEs (1-3). However, the HPM can suffer from round-off errors in some cases. Therefore, the focus of this study is to propose a hybrid approach that combines the HPM with the Laplace Transform Method (LTM) to improve the accuracy of the numerical solutions. Various methods have been developed for solving PDEs, such as Finite Difference Method (FDM), Finite Element Method (FEM), Differential Transform Method (4) and Spectral Method. However, these methods often require intricate discretization techniques that can lead to inaccurate solutions (5). The HPM, developed by He, offers a promising alternative approach that avoids these issues by finding a solution in the form of a series of homotopy functions that the Taylor series can rapidly and effortlessly approximate. The HPM has recently been used to solve several PDEs (6-8).

Although highly effective, the HPM has some limitations, particularly in dealing with nonlinear terms. To address this issue, a new method called the He-Laplace method has been introduced, which combines the HPM with He's polynomials to handle nonlinearities in PDEs (8, 9). The proposed method has been applied to linear and nonlinear PDEs, exhibiting superior accuracy and efficiency to traditional numerical methods (10, 11). Researchers have favored the homotopy perturbation method (HPM) as an analytical technique to solve a range of Heat transport and fluid mechanics differential equations in recent studies (12-17). In a recent paper by Jalili et al. (18), the HPM and AGM method was used to resolve thermal problems for micro-polar nanofluids. Alternatively, Jalili et al. (19) applied the HPM method to tackle problems related to the movement of a plate submerged in a Newtonian fluid examined through Caputo fractional differential equations. Jalili et al. (20) also demonstrated a novel fractional analytical approach to solve time-space fractional equations from oil pollution analysis using HPM.

Python programming has become a common tool in scientific computing due to its simplicity, flexibility, and efficient numerical package libraries, such as NumPy, SciPy, and Matplotlib. The availability of open-source libraries for Python has enabled the development of fast and efficient algorithms for solving PDEs using numerical methods (21-23). The HPM has recently been used with other methods to improve its accuracy and efficiency. The HPM-Laplace method, which combines the HPM with the Laplace Transform Method (LTM) to increase the accuracy of the numerical solutions, has been proposed. The method has been used to solve various linear and nonlinear PDEs, exhibiting superior accuracy and efficiency compared to traditional numerical methods (24, 25). The HPM-Laplace method could be implemented in Python to solve PDE problems with high accuracy and efficiency. Another hybrid approach involving the HPM is He's polynomial-based decomposition method, which uses He's polynomials to decompose the solution of a PDE into several parts, for which each part can be solved by a different method (26). The decomposition method involves the solution of smaller sub-problems, which can easily be implemented using Python functions. Also, much research has been done on the analytical and Python approach (27-34).

In regards to the applications of solving PDEs, the following articles shed light on diverse areas of research, including the dynamics of modified Peyrard-Bishop DNA models in bio-fluids, the analysis of protein oscillations through Langevin and Fokker-Planck equations, and the optimization of heat exchanger design for enhanced heat transfer rates and temperature distribution. Lap-Arparat and Tuchinda (35) explored the behavior of the modified Peyrard-Bishop DNA model within a thermostat as a bio-fluid. They investigate the soliton-like solutions in this system and discuss their implications for understanding the dynamics of DNA molecules. Baughman and Sharma (36) focused on the entropic analysis of protein oscillations. They employed Langevin and Fokker-Planck equations to study the dynamics of proteins and analyzed the role of entropy in protein oscillations. Their findings contribute to a deeper understanding of the complex behavior of proteins. In the study conducted by Sutantyo et al. (37), they have investigated the impact of heat exchanger design on heat transfer rate and temperature distribution. They analyzed different heat exchanger designs and evaluate their heat transfer efficiency and temperature distribution performance. The results of their study provided valuable insights for optimizing heat exchanger design to enhance heat transfer processes.

In the realm of solving partial differential equations (PDEs), significant difficulties and challenges hinder the attainment of accurate and reliable solutions. One of the primary obstacles lies in effectively handling equations with nonlinear components, which often lead to complex mathematical formulations. To overcome this challenge, we have developed a unique hybrid technique, HPM-LTM, which combines the advantages of the Homotopy Perturbation technique (HPM) and the Laplace Transform Method (LTM). By employing He's polynomials, our approach successfully tackles nonlinear terms, enabling the efficient and accurate solution of PDEs. This original achievement allows us to bypass the limitations imposed by conventional approaches that are often plagued by numerical errors and restrictive assumptions. By implementing the HPM-LTM method, we have achieved remarkable accuracy and reliability in solving PDEs without discretization, thereby eliminating potential round-off errors. We develop Python functions to implement the proposed method and use the numerical package libraries to solve five PDE examples in various engineering fields. The examples have been chosen to demonstrate the applicability and flexibility of our method in various domains. Specifically, the examples include linear PDEs, such as the heat and wave equation, and nonlinear PDEs, such as the Burgers' Equation, the MHD flow equation, and the Klein-Gordon equation. By including a diverse set of examples, ranging from linear to nonlinear PDEs, we aim to showcase the effectiveness of our method across various engineering domains. Adopting this approach allows us to assess the method's performance under different types of PDEs, capturing a wider scope of real-world scenarios.

The article is structured as follows: Section 1 provides an introduction, offering an overview of the problem at hand and a review of the relevant literature. Section 2 focuses on the mathematical formulation, explaining the fundamental concepts of the HPM and LTM methods. Section 3, titled methodology, delves into the basic idea of the HPM-LTM method and provides a detailed explanation of SymPy and differential equations. Section 4 presents five examples and their solutions, showcasing the results obtained for both linear and nonlinear partial differential equations. Finally, section 5 concludes the article, summarizing the findings and highlighting the key advantages of the HPM-LTM method.

2. MATHEMATICAL FORMULATION

Numerous scientific investigations rely heavily on mathematical formulations because they give researchers a mechanism to properly define and measure the physical processes under investigation. We shall examine how our current problem is mathematically formulated in this part. We shall start by describing the issue statement and stating our goals. The key mathematical theories and equations required to resolve the issue will then be covered. We aim to give a clear and concise grasp of the underlying ideas and techniques required to solve by decomposing the issue into its constituent mathematical parts.

2. 1. The Basic Idea of HPM

We examined the subsequent equation to elucidate the underlying concepts:

$$w(k) - h(j) = 0, \quad j\epsilon\Gamma \tag{1}$$

With the B.Cs. of:

$$B\left(k, \frac{\partial k}{\partial n}\right) = 0, j\epsilon\delta \tag{2}$$

where B is a boundary operator, δ is a domain (Γ) boundary, h (j) is a known analytic function, and w (k) is a differential operator. Therefore, Equation 1 can be rewritten as Equation 3, $L(k)$ is linear, and $N(k)$ is nonlinear.

$$L(k) + N(k) - h(j) = 0,\ j\epsilon\Gamma \tag{3}$$

The HPM configuration looks like Equation 4:

$$H(q,p) = (1-p)[L(q) - L(k_0)] + p[w(q) - h(j)] = 0 \tag{4}$$

$$q(j,p)\colon \Gamma \times [0,1] \to R \tag{5}$$

The initial approximation in Equation 5 that satisfies the B.Cs. is k_0, and p is an embedded parameter that falls between [0,1]. The answer to Equation 4 can be expressed as follows in powers of p:

$$q = q_0 + pq_1 + p^2 q_2 + \cdots\cdots\cdots \tag{6}$$

And the best guess for the answer is:

$$k = Lim_{p\to 1} q = q_0 + q_1 + q_2 + \cdots\cdots\cdot \tag{7}$$

2. 2. The Basic Idea of LTM

Operators, also known as transformations, turn one function into another. Although transformations alter the original function, they can maintain or change the independent variable. For example, when a function $f(t)$ is multiplied by a number n using the operator O, the resulting function $g(t)$ keeps the original function's independent variable. Formally, such an operator may be stated as:

$$O[f(t)] = n \cdot f(t) = g(t) \tag{8}$$

The inverse Laplace transform acts on $F(s)$ to get $f(t)$, whereas the direct Laplace transform takes a function $f(t)$ that relies on time t and converts it into a different function $F(s)$ that depends on the complex variable s. The Laplace transform is described mathematically as

$$F(s) = \mathcal{L}[f(t)] = \int_0^\infty f(t) \cdot e^{-s \cdot t} dt \tag{9}$$

For the transformation to occur, the integral defining the Laplace transform must converge, which necessitates that the initial function $f(t)$ meet the following prerequisites:

1) If the interval can be divided into a finite number of nonintersecting intervals, then $f(t)$ must be piecewise continuous across the interval $0 < t < \infty$, which means it must be continuous on any subinterval of that interval, and it must have finite bounds at the endpoints of each subinterval.
2) Given that $f(t)$ must be of exponential order, it must meet the equality for real constant values σ

that is greater than a critical value, σ_c, often known as the abscissa of convergence.

Formally, the inverse Laplace transform, which is often needed to return the original (generally unknown) time-domain function $f(t)$, is calculated as:

$$f(t) = \mathcal{L}^{-1}[F(s)] \tag{10}$$

3. METHODOLOGY

3. 1. The Basic Idea of HPM-LTM

Up to step (4), HPM-LTM follows the same stages as HPM. Then, we apply the Laplace transform to both sides of the homotopy Equation 4:

$$\mathfrak{s}\{L(K) - L(k_0) + p[L(k_0) + N(K) - h(j)]\} = 0 \tag{11}$$

The possibility exists that HPM-LTM could replace (k_0), with any arbitrary function $z(x)$, provided that this function is defined by certain unspecified parameters A, B, C, etc. These parameters would need to be appropriately determined using the flexibility afforded by the homotopy formulation. Then, we have:

$$\begin{aligned} s^n\mathfrak{I}\{K\} - s^{n-1}K(0) - s^{n-2}K'(0) - \cdots - K^{(n-1)}(0) \\ = \mathfrak{I}\{L(k_0) - pL(k_0) + p[-N(K) + h(j)]\} \end{aligned} \tag{12}$$

Or

$$\begin{aligned} \mathfrak{I}(K) = {} & \left(\tfrac{1}{s^n}\right)\{s^{n-1}K(0) + s^{n-2}K'(0) + \cdots + K^{(n-1)}(0)\} \\ & + \left(\tfrac{1}{s^n}\right)\mathfrak{I}\{L(k_0) - pL(k_0) + p[-N(K) + h(j)]\}, \end{aligned} \tag{13}$$

When we apply the inverse Laplace transform to Equation 13, we get:

$$\begin{aligned} K = {} & \mathfrak{I}^{-1}\Big\{\left(\tfrac{1}{s^n}\right)\{s^{n-1}K(0) + s^{n-2}K'(0) + \cdots + K^{(n-1)}(0)\} \\ & + \left(\tfrac{1}{s^n}\right)\mathfrak{I}\{L(k_0) - pL(k_0) + p[-N(K) + h(j)]\}\Big\}. \end{aligned} \tag{14}$$

Assuming that a power series of p may be used to describe the answers to Equation 3:

$$K = \sum_{n=0}^{\infty} p^n q_n \tag{15}$$

then we obtain by replacing Equation 15 with Equation 14.

$$\sum_{n=0}^{\infty} p^n q_n = \mathfrak{I}^{-1}\left\{\begin{array}{l} \left(\frac{1}{s^n}\right)\{s^{n-1}K(0) + s^{n-2}K'(0) + \cdots + K^{(n-1)}(0)\} \\ + \left(\frac{1}{s^n}\right)\mathfrak{I}\{L(k_0) - pL(k_0) + p[-N(\sum_{n=0}^{\infty} p^n q_n) + h(j)]\} \end{array}\right\} \tag{16}$$

Comparing p coefficients of equal power results in

$$\begin{aligned} & p^0: q_0 = \mathfrak{I}^{-1}\left\{\left(\tfrac{1}{s^n}\right)\left(s^{n-1}K(0) + s^{n-2}K'(0) + \cdots + K^{(n-1)}(0)\right) + \mathfrak{I}\{L(k_0)\}\right)\Big\}, \\ & p^1: q_1 = \mathfrak{I}^{-1}\left\{\left(\tfrac{1}{s^n}\right)(\mathfrak{I}\{-N_0(q_0) - L(k_0) + h(rj)\})\right\}, \\ & p^2: q_2 = \mathfrak{I}^{-1}\left\{\left(\tfrac{1}{s^n}\right)\mathfrak{I}\{-N_1(q_0, q_1)\}\right\} \\ & p^3: q_3 = \mathfrak{I}^{-1}\left\{\left(\tfrac{1}{s^n}\right)\mathfrak{I}\{-N_2(q_0, q_1, q_2)\}\right\}, \\ & p^j: q_j = \mathfrak{I}^{-1}\left\{\left(\tfrac{1}{s^n}\right)\mathfrak{I}\{-N_{j-1}(q_0, q_1, q_2, \ldots, q_{j-1})\}\right\} \end{aligned} \tag{17}$$

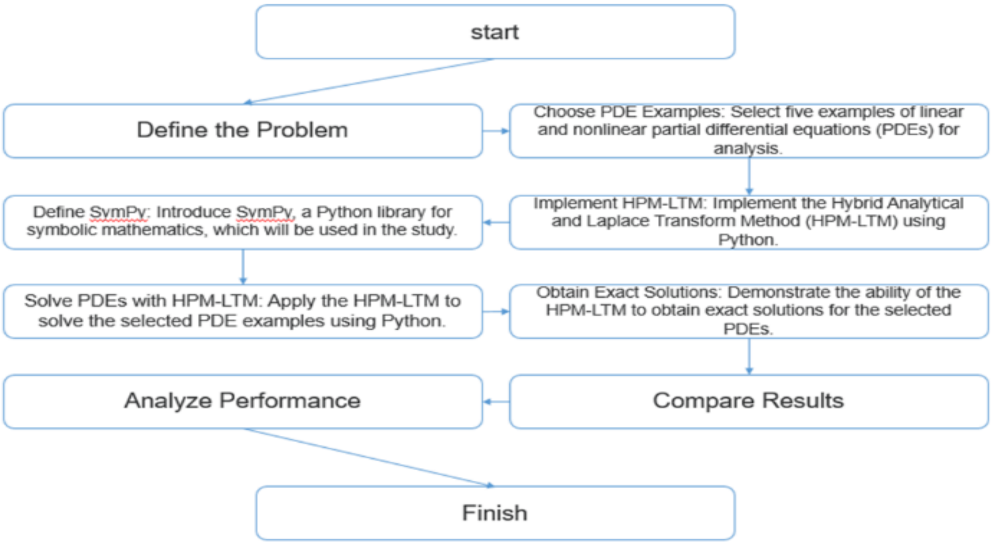
Assuming that $K(0) = k_0 = \alpha_0, K'(0) = \alpha_1, \ldots, K^{n-1}(0) = \alpha_{n-1}$; the exact solution may be obtained as follows:

$$k = \lim_{p\to 1} K = q_0 + q_1 + q_2 + \cdots \tag{18}$$

The notation presented above indicates the potential for extending the nonlinear operator N as a formal series with respect to the embedding parameter p.

3. 2. SymPy and Differential Equations

In our study, we harnessed the capabilities of Python in conjunction with the SymPy module, a powerful tool for conducting symbolic computations. Particularly in algebra, discrete mathematics, and calculus, we found SymPy exceptionally useful, primarily due to its proficiency in formatting and presenting outcomes in LaTeX format. We employed a multi-step procedure to address differential equations using SymPy. Commencing with importing the SymPy library and associated symbols, we defined the dependent variable as a function of the independent counterpart. The incorporation of boundary or initial conditions was achieved by utilizing the '.subs' method. We then embarked on the creation of the differential equation itself. Subsequently, the crux of our endeavor lay in solving this intricate differential equation by applying the 'dsolve' function. To streamline our efforts, we availed ourselves of the 'from sympy import *' and 'from sympy.abc import *' directives facilitated the importation of requisite modules and symbols. The complex differential equations central to our research were successfully unraveled, all thanks to the robust support provided by SymPy. This, in turn, furnished us with profound insights that greatly enriched our understanding. To clarify the methodology, we provide the following flowchart:



4. ILLUSTRATIVE EXAMPLES

We have applied the HPM-LTM method to demonstrate its effectiveness in two linear and three nonlinear partial differential equations.

Example 4.1. The homogeneous linear PDE (26, 36, 38) is as follows:

$$\frac{\partial y}{\partial t}+\frac{\partial y}{\partial x}-\frac{\partial^2 y}{\partial x^2}=0 \tag{19}$$

with the following B.Cs.:

$$y(x,0)=e^x-x,\ y(0,t)=1+t,\ \frac{\partial y}{\partial x}(1,t)=e-1 \tag{20}$$

Applying the procedure above under the presumption that existed initially, we have:

$$y(x,s)=\frac{e^x-x}{s}-\frac{1}{s}L\left[\frac{\partial y}{\partial x}-\frac{\partial^2 y}{\partial x^2}\right] \tag{21}$$

The Laplace transform's inverse suggests that:

$$y(x,t)=e^x-x-L^{-1}\left[\frac{1}{s}L\left[\frac{\partial y}{\partial x}-\frac{\partial^2 y}{\partial x^2}\right]\right] \tag{22}$$

Now, we apply the HPM; we have

$$\sum_{n=0}^{\infty} p^n y_n(x,t)=e^x-x-p\left(L^{-1}\left[\frac{1}{s}L\left[\frac{\partial y}{\partial x}-\frac{\partial^2 y}{\partial x^2}\right]\right]\right) \tag{23}$$

The coefficient of similar powers of p is compared, and we have:

$$\begin{aligned} &p^0: y_0(x,t)=e^x-x\\ &p^1: y_1(x,t)=-L^{-1}\left[\frac{1}{s}L\left[\frac{\partial y_0}{\partial x}-\frac{\partial^2 y_0}{\partial x^2}\right]\right]=t\\ &p^2: y_2(x,t)=-L^{-1}\left[\frac{1}{s}L\left[\frac{\partial y_1}{\partial x}-\frac{\partial^2 y}{\partial x^2}\right]\right]=0 \end{aligned} \tag{24}$$

$$\begin{aligned} &p^3: y_3(x,t)=0,\\ &p^4: y_4(x,t)=0, \end{aligned} \tag{25}$$

so the solution $y(x,t)$ is given by

$$\begin{aligned} y(x,t)\ &=e^x-x+t+0+0\cdots\\ &=e^x-x+t \end{aligned} \tag{26}$$

We begin by noting that Equation 26 represents the exact solution of the problem, as reported in literature (26, 36, 38). Different approximate solutions are obtained in some experiments using HPM (36, 38). By applying HPM-LTM, we obtain exact solutions.

The problem was solved using Python, leveraging the capabilities of the symbolic mathematics library Sympy. The results obtained from the solution are presented in Figures 1 and 2. Figure 1 depicts the exact solution of $y(t)=e^x-x+t$ for various values of t. It is evident that as the value of x increases, the solution of $y(t)$ exhibits exponential growth. The different lines in the figure correspond to different values of t in the exact solution of the equation. This visualization provides valuable insights into the behavior of the solution as it varies with different input parameters. Figure 2 presents a 2D contour plot that offers a clearer visualization of the exponential increase in t as x increases. The contour lines represent equal $y(t)$ values for different combinations of x and t. By examining this contour plot, we understand how the solution behaves across the entire parameter space. The contour lines help identify regions where the solution exhibits significant changes or remains relatively constant. Both Figures 1 and 2 collectively demonstrate the exponential behavior of the exact solution of the given equation. These visual representations provide valuable insights and enhance our understanding of the problem. Employing the

Matplotlib library in Python facilitated the creation of these plots, enabling the effective communication of the obtained results.

The results obtained through the utilization of Python, Sympy, and Matplotlib highlight the exponential growth of the exact solution of $y(t) = e^x - x + t$. The figures presented in this section provide a comprehensive visualization of this behavior, showcasing the impact of varying input parameters on the solution. These findings contribute to a deeper understanding of the problem and validate the effectiveness of this study's hybrid analytical and Laplace transform methods.

Example 4.2. Take into account the homogeneous linear PDE (Klein-Gordon equation) shown below (26, 36, 38):

$$\frac{\partial^2 y}{\partial t^2} + y - \frac{\partial^2 y}{\partial x^2} = 0 \qquad (27)$$

with the following B.Cs.:

$$y(x,0) = e^{-x} + x,\ \frac{\partial y}{\partial t}(x,0) = 0 \qquad (28)$$

Applying the procedure above by the I.Cs., we have:

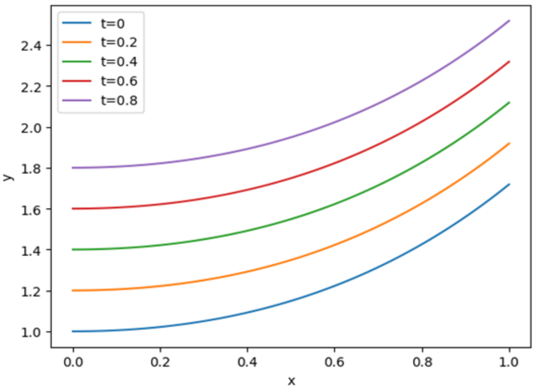


Figure 1. Exact solution of example 4.1 using HPM-LTM method for varying values of t, plotted in Python using Sympy and Matplotlib for different values of t

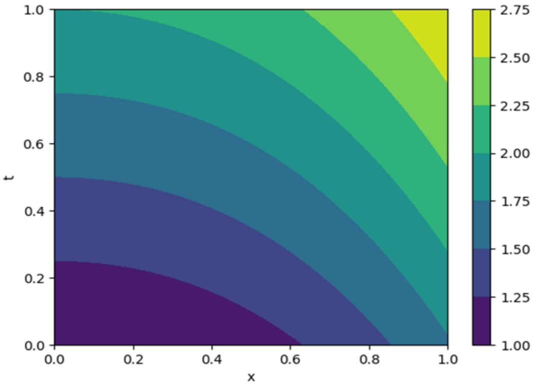


Figure 2. 2D contour plot of the exact solution of example 4.1 using the HPM-LTM method, plotted in Python using Sympy and Matplotlib

$$y(x,s) = \frac{e^{-x}+x}{s} - \frac{1}{s^2} L\left[y - \frac{\partial^2 y}{\partial x^2}\right] \qquad (29)$$

The Laplace transform's inverse suggests that:

$$y(x,t) = e^{-x} + x - L^{-1}\left[\frac{1}{s^2} L\left[y - \frac{\partial^2 y}{\partial x^2}\right]\right] \qquad (30)$$

Now, we apply the HPM; we have

$$\sum_{n=0}^{\infty} p^n y_n(x,t) = e^{-x} + x - p\left(L^{-1}\left[\frac{1}{s^2} L\left[y - \frac{\partial^2 y}{\partial x^2}\right]\right]\right). \qquad (31)$$

The coefficient of similar powers of p is compared, and we have:

$$\begin{aligned} p^0 &: y_0(x,t) = e^{-x} + x \\ p^1 &: y_1(x,t) = -L^{-1}\left[\frac{1}{s^2} L\left[y_0 - \frac{\partial^2 y_0}{\partial x^2}\right]\right] = \frac{-xt^2}{2!}, \\ p^2 &: y_2(x,t) = -L^{-1}\left[\frac{1}{s^2} L\left[y_1 - \frac{\partial^2 y_1}{\partial x^2}\right]\right] = \frac{xt^4}{4!}. \end{aligned} \qquad (32)$$

$$\begin{aligned} p^3 &: y_3(x,t) = \frac{-xt^6}{6!} \\ p^4 &: y_4(x,t) = \frac{xt^8}{8!} \\ p^n &: y_n(x,t) = \frac{(-1)^n xt^{2n}}{2n!} \end{aligned} \qquad (33)$$

so that the solution $y(x,t)$ is given by:

$$\begin{aligned} y(x,t) &= y_0 + y_1 + y_2 + y_3 + \cdots \\ &= e^{-x} + x - \frac{xt^2}{2!} + \frac{xt^4}{4!} - \frac{xt^6}{6!} + \cdots + \frac{(-1)^n xt^{2n}}{2n!} \\ &= e^{-x} + x\left(1 - \frac{t^2}{2!} + \frac{t^4}{4!} - \cdots + \frac{(-1)^n t^{2n}}{2n!} + \cdots\right) \\ &= e^{-x} + xCos\,(t), \end{aligned} \qquad (34)$$

Equation 34 is the exact solution to the problem (26, 36, 38), which is a homogeneous linear PDE (Klein-Gordon equation), depicted in Figures 3 and 4, using the HPM-LTM method in Python (26, 36, 38). Different approximate solutions are obtained in some experiments using HPM (36, 38). By applying HPM-LTM, we obtain exact solutions.

Figure 3 specifically focuses on the behavior of the solution over time or for different boundary conditions. It presents a line graph illustrating the exact solution of the equation for different values of t. The graph clearly demonstrates how the solution evolves over time, providing insights into its temporal dynamics. As t increases, the corresponding value of y decreases. This decreasing trend suggests that the amplitude of the solution diminishes as time progresses. Notably, a particularly noticeable drop in the solution occurs at $t = 0.8$. This observation indicates that the solution becomes less pronounced as time advances. By visualizing the solution's behavior for various values of t, Figure 3 effectively showcases the temporal evolution and sensitivity of the solution to different time-dependent factors. This plot provides valuable insights into how the solution changes over time and in response to varying boundary conditions. Figure 4, on the other hand, offers a 2D contour plot that represents different solution levels based on y and x values. This contour plot allows us to identify variations in the solution across the $y-x$ parameter space, enabling the identification of patterns, trends, and critical points. We can observe regions of equal solution values by examining the contour lines in Figure 4. This visualization helps us identify areas where the solution exhibits significant changes or remains relatively constant. Consequently, we can pinpoint critical points, such as regions with steep gradients or areas where the solution transitions rapidly. The 2D contour plot presented in Figure 4 enhances our understanding of the solution by providing a comprehensive overview of its variations across different y and x values.

Figures 3 and 4 provide a detailed analysis of the exact solution of the homogeneous linear PDE obtained using the HPM-LTM method in Python. Figure 3's line graph demonstrates the solution's temporal behavior, showcasing a decreasing trend in the solution as t increases. On the other hand, Figure 4's 2D contour plot offers a comprehensive visualization of the solution's variations across the $y-x$ parameter space, allowing us to identify patterns, trends, and critical points. Both figures contribute to a deeper understanding of the solution and further validate the effectiveness of the HPM-LTM method for solving the considered homogeneous linear PDE.

Example 4.3. Take into account the homogeneous nonlinear PDE (Burger equation) shown below (26, 36, 38):

$$\frac{\partial y}{\partial t} - y\frac{\partial y}{\partial x} - \frac{\partial^2 y}{\partial x^2} = 0 \tag{35}$$

with the following B.Cs.:

$$y(x,0) = 1-x,\ y(0,t) = \frac{1}{(1+t)},\ y(1,t) = 0 \tag{36}$$

Applying the procedure above while keeping in mind the starting state, we have:

$$y(x,s) = \frac{1-x}{s} + \frac{1}{s}L\left[\frac{\partial^2 y}{\partial x^2} + y\frac{\partial y}{\partial x}\right] \tag{37}$$

The Laplace transform's inverse suggests that:

$$y(x,t) = 1-x+L^{-1}\left[\frac{1}{s}L\left[\frac{\partial^2 y}{\partial x^2} + y\frac{\partial y}{\partial x}\right]\right] \tag{38}$$

Now, we apply the HPM; we have

$$\sum_{n=0}^{\infty} p^n y_n(x,t) = 1-x+p\left(L^{-1}\left[\frac{1}{s}\left\{L\left[\frac{\partial^2 y}{\partial x^2}\right] + L\left[\sum_{n=0}^{\infty} p^n H_n(y)\right]\right\}\right]\right), \tag{39}$$

where $H_n(y)$ are He's polynomials. The first few components of He's polynomials are:

$$\begin{aligned} H_0(y) &= y_0\frac{\partial y_0}{\partial x} = -(1-x) \\ H_1(y) &= y_0\frac{\partial y_1}{\partial x} + y_1\frac{\partial y_0}{\partial x} = 2(1-x)t \\ H_2(y) &= y_0\frac{\partial y_2}{\partial x} + y_1\frac{\partial y_1}{\partial x} + y_2\frac{\partial y_0}{\partial x} = -3(1-x)t^2 \end{aligned} \tag{40}$$

The coefficient of similar powers of p is compared, and we have:

$$\begin{aligned} p^0&: y_0(x,t) = 1-x \\ p^1&: y_1(x,t) = L^{-1}\left[\frac{1}{s}\left\{L\left[\frac{\partial^2 y_0}{\partial x^2}\right] + L[H_0(y)]\right\}\right] = -(1-x)t \\ p^2&: y_2(x,t) = L^{-1}\left[\frac{1}{s}\left\{L\left[\frac{\partial^2 y_1}{\partial x^2}\right] + L[H_1(y)]\right\}\right] = (1-x)t^2 \end{aligned} \tag{41}$$

$$\begin{aligned} p^3&: y_3(x,t) = -(1-x)t^3 \\ p^4&: y_4(x,t) = (1-x)t^4 \end{aligned} \tag{42}$$

so that the solution $y(x,t)$ is given by:

$$\begin{aligned} y(x,t) &= y_0 + y_1 + y_2 + y_3 + \cdots \\ &= (1-x) - (1-x)t + (1-x)t^2 - (1-x)t^3 + \cdots \\ &= (1-x)[1 - t + t^2 - t^3 + t^4 - \cdots] \\ &= (1-x)(1+t)^{-1} = \frac{(1-x)}{(1+t)}, \end{aligned} \tag{43}$$

which is the exact solution to the problem (26, 36, 38). In some experiments, different approximate solutions are obtained by using HPM (36, 38). By applying HPM-LTM, we obtain exact solutions. Figure 5 presents a line graph that showcases the behavior of the solution for various values of t $(0, 0.2, 0.4, 0.6, 0.8)$ in relation to the value of y. The graph provides valuable insights into the dynamics of the solution and its sensitivity to changes in t.

The observed trend in Figure 5 indicates that the solution curve is decreasing, with y decreasing as t increases. This behavior implies that the amplitude of the solution diminishes as time progresses. This decreasing trend is evident in the lines representing $t = 0.2, t = 0.4, t = 0.6$, and $t = 0.8$. However, an interesting anomaly is observed in the curve for t = 0, where the solution starts with $y = 1$ and decreases as t increases. This reversed trend in the line plot suggests a unique behavior or specific conditions that influence the solution's initial behavior. Another noteworthy observation in Figure 5 is the significant increase in the solution's value at $t = 0.8$. This sharp increase indicates a sudden change or a critical point in the solution's behavior. It suggests that the solution becomes more pronounced or exhibits a higher amplitude at this specific time point.

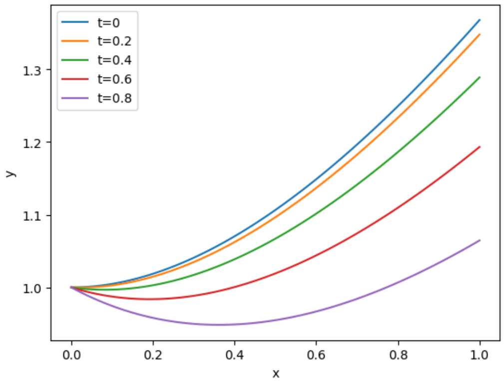


Figure 3. Exact solution of example 4.2 using HPM-LTM method for varying values of t, plotted in Python using Sympy and Matplotlib for different values of t

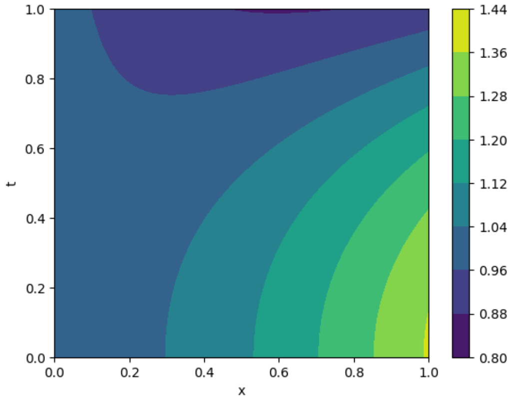


Figure 4. 2D contour plot of the exact solution of example 4.2 using the HPM-LTM method, plotted in Python using Sympy and Matplotlib

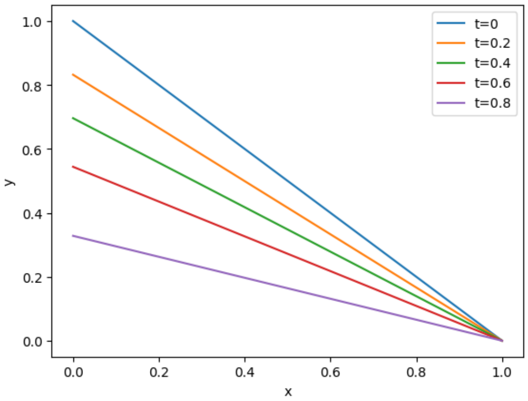


Figure 5. Exact solution of example 4.3 using HPM-LTM method for varying values of t, plotted in Python using Sympy and Matplotlib for different values of t

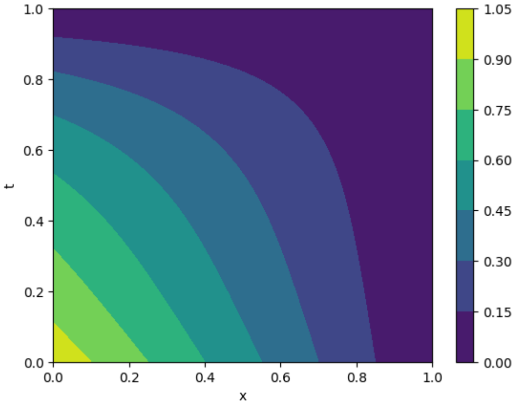


Figure 6. 2D contour plot of the exact solution of example 4.3 using the HPM-LTM method, plotted in Python using Sympy and Matplotlib

Additionally, all the lines in the graph converge at $x = 1$, indicating that the behavior of the solution remains consistent across different values of t as it approaches this point. This convergence implies that the solution reaches a stable state or exhibits similar characteristics near $x = 1$, regardless of the specific value of t. It suggests that the solution may have a specific behavior or property that is invariant with respect to changes in t near this point.

Figure 6 appears to be a 2D contour plot that provides further insights into the behavior of the solution for different values of t and x. The contour lines on the plot represent different levels of the solution and allow us to visualize how the values change across various regions of the x and t axes. By examining the contour lines in Figure 6, we can identify variations in the solution across the $x - t$ parameter space. This visualization lets us identify regions with similar solution values and observe patterns, trends, and critical points. The contour plot helps us understand how the solution changes across different values of t and x. It provides a comprehensive overview of the solution's behavior and allows us to identify areas where the solution exhibits significant changes or remains relatively constant. By analyzing the contour lines, we can identify regions with steep gradients or areas where the solution transitions rapidly. This information is crucial for understanding the behavior and characteristics of the solution in different parts of the x-t parameter space.

Figure 5 illustrates the behavior of the solution for various values of t, showing a decreasing trend overall with an interesting reversal in the curve for $t = 0$. Figure 6, on the other hand, provides a 2D contour plot that offers insights into the behavior of the solution across different values of t and x. The contour lines help to identify the solution's variations, patterns, and critical points. These visualizations enhance our understanding of the physical behavior and properties of the solution in the context of the given problem.

Example 4.4. Take into account the aforementioned homogeneous nonlinear PDE (26, 36, 38):

$$\frac{\partial y}{\partial t}-y-y\frac{\partial^2 y}{\partial x^2}-\left(\frac{\partial y}{\partial x}\right)^2=0 \tag{44}$$

with the following B.Cs.:

$$y(x,0)=\sqrt{x},\ y(0,t)=0,\ y(1,t)=e^t. \tag{45}$$

Applying the procedure above while keeping in mind the starting state, we have:

$$y(x,s)=\frac{\sqrt{x}}{s}+\frac{1}{s}L\left[y+y\frac{\partial^2 y}{\partial x^2}+\left(\frac{\partial y}{\partial x}\right)^2\right] \tag{46}$$

The Laplace transform's inverse suggests that:

$$y(x,t)=\sqrt{x}+L^{-1}\left[\frac{1}{s}L\left[y+y\frac{\partial^2 y}{\partial x^2}+\left(\frac{\partial y}{\partial x}\right)^2\right]\right] \tag{47}$$

Now, we apply the HPM; we have

$$\sum_{n=0}^{\infty} p^n y_n(x,t)=\sqrt{x}+p\left(L^{-1}\left[\frac{1}{s}\{L[y]+L[\textstyle\sum_{n=0}^{\infty} p^n H_n(y)]\}\right]\right) \tag{48}$$

where $H_n(y)$ are He's polynomials. The first few components of He's polynomials are given by:

$$\begin{aligned}
H_0(y)&=y_0\frac{\partial^2 y_0}{\partial x^2}+\left(\frac{\partial y_0}{\partial x}\right)^2=0\\
H_1(y)&=y_0\frac{\partial^2 y_1}{\partial x^2}+y_1\frac{\partial^2 y_0}{\partial x^2}+2\frac{\partial y_0}{\partial x}\frac{\partial y_1}{\partial x}=0\\
H_2(y)&=y_0\frac{\partial^2 y_2}{\partial x^2}+y_1\frac{\partial^2 y_1}{\partial x^2}+y_2\frac{\partial^2 y_0}{\partial x^2}+\left(\frac{\partial y_1}{\partial x}\right)^2+2\frac{\partial y_0}{\partial x}\frac{\partial y_2}{\partial x}=0
\end{aligned} \tag{49}$$

The coefficient of similar powers of p is compared, and we have:

$$\begin{aligned}
&p^0: y_0(x,t)=\sqrt{x}\\
&p^1: y_1(x,t)=L^{-1}\left[\frac{1}{s}\{L[y_0]+L[H_0(y)]\}\right]=\sqrt{x}t\\
&p^2: y_2(x,t)=L^{-1}\left[\frac{1}{s}\{L[y_1]+L[H_1(y)]\}\right]=\frac{\sqrt{x}t^2}{2!}.
\end{aligned} \tag{50}$$

$$\begin{aligned}
&p^3: y_3(x,t)=\frac{\sqrt{x}t^3}{3!},\\
&p^4: y_4(x,t)=\frac{\sqrt{x}t^4}{4!},
\end{aligned} \tag{51}$$

so the solution $y(x,t)$ is given by:

$$\begin{aligned}
y(x,t)\quad &=y_0+y_1+y_2+y_3+\cdots\\
&=\sqrt{x}+\frac{\sqrt{x}t}{1!}+\frac{\sqrt{x}t^2}{2!}+\frac{\sqrt{x}t^3}{3!}+\cdots\\
&=\sqrt{x}\left(1+\frac{t}{1!}+\frac{t^2}{2!}+\frac{t^3}{3!}+\cdots+\frac{t^n}{n!}+\cdots\right)\\
&=\sqrt{x}e^t,
\end{aligned} \tag{52}$$

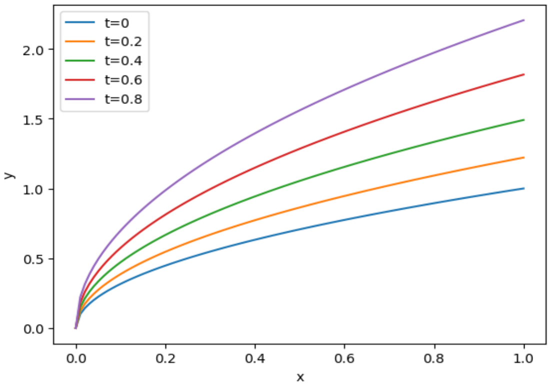


Figure 7. Exact solution of example 4.4 using HPM-LTM method for varying values of t, plotted in Python using Sympy and Matplotlib for different values of t

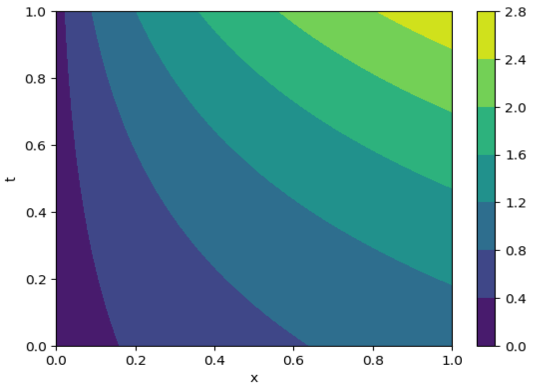


Figure 8. 2D contour plot of the exact solution of example 4.4 using the HPM-LTM method, plotted in Python using Sympy and Matplotlib

which is the exact solution to the problem (26, 36, 38). In some experiments, different approximate solutions are obtained by using HPM (36, 38). By applying HPM-LTM, we obtain exact solutions.

Example 4.5. Take into account the subsequent non-homogeneous nonlinear PDE (26, 36, 38):

$$\frac{\partial^2 y}{\partial t^2}+\frac{\partial^2 y}{\partial x^2}+\left(\frac{\partial y}{\partial x}\right)^2=2x+t^4 \tag{53}$$

with the following B.Cs.:

$$y(x,0)=0,\ \frac{\partial y}{\partial t}(x,0)=a,\ y(0,t)=at,\ \frac{\partial y}{\partial x}(0,t)=t^2 \tag{54}$$

By applying the HPM subject to the I.Cs., we have:

$$\begin{aligned}
y(x,s)\quad &=\frac{a}{s^2}-\frac{1}{s^2}L\left[\frac{\partial^2 y}{\partial x^2}+\left(\frac{\partial y}{\partial x}\right)^2\right]+\frac{1}{s^2}L[2x+t^4]\\
&=\frac{a}{s^2}+\frac{2x}{s^3}+\frac{4!}{s^7}-\frac{1}{s^2}L\left[\frac{\partial^2 y}{\partial x^2}+\left(\frac{\partial y}{\partial x}\right)^2\right].
\end{aligned} \tag{55}$$

The Laplace transform's inverse suggests that:

$$y(x,t) = at + xt^2 + \frac{t^6}{30} - L^{-1}\left[\frac{1}{s^2} L\left[\frac{\partial^2 y}{\partial x^2} + \left(\frac{\partial y}{\partial x}\right)^2\right]\right]. \qquad (56)$$

Now, we apply the HPM, we have:

$$\sum_{n=0}^{\infty} p^n y_n(x,t) = at + xt^2 + \frac{t^6}{30} - p\left(L^{-1}\left[\frac{1}{s^2}\left\{L\left[\frac{\partial^2 y}{\partial x^2}\right] + L\left[\sum_{n=0}^{\infty} p^n H_n(y)\right]\right\}\right]\right) \qquad (57)$$

where $H_n(y)$ are He's polynomials. The first few components of He's polynomials are given by:

$$\begin{aligned} H_0(y) &= \left(\frac{\partial y_0}{\partial x}\right)^2 = t^4, \\ H_1(y) &= 2\left(\frac{\partial y_0}{\partial x}\right) \times \left(\frac{\partial y_1}{\partial x}\right) = 0, \\ H_2(y) &= \left(\frac{\partial y_1}{\partial x}\right)^2 + 2\frac{\partial y_0}{\partial x}\frac{\partial y_2}{\partial x} = 0, \end{aligned} \qquad (58)$$

The coefficient of similar powers of p is compared, and we have:

$$\begin{aligned} p^0 &: y_0(x,t) = at + xt^2 + \frac{t^6}{30}, \\ p^1 &: y_1(x,t) = -L^{-1}\left[\frac{1}{s^2}\left\{L\left[\frac{\partial^2 y_0}{\partial x^2}\right] + L[H_0(y)]\right\}\right] = -\frac{t^6}{30}, \\ p^2 &: y_2(x,t) = -L^{-1}\left[\frac{1}{s^2}\left\{L\left[\frac{\partial^2 y_1}{\partial x^2}\right] + L[H_1(y)]\right\}\right] = 0. \end{aligned} \qquad (59)$$

$$\begin{aligned} p^3 &: y_3(x,t) = 0, \\ p^4 &: y_4(x,t) = 0, \end{aligned} \qquad (60)$$

so that the solution $y(x,t)$ is given by:

$$\begin{aligned} y(x,t) &= y_0 + y_1 + y_2 + y_3 + \cdots \\ &= at + xt^2 + \frac{t^6}{30} - \frac{t^6}{30} + 0 + 0 + 0 + \cdots \\ &= at + xt^2 \end{aligned} \qquad (61)$$

which is the exact solution to the problem (26, 36, 38). In some experiments, different approximate solutions are obtained by using HPM (36, 38). By applying HPM-LTM, we obtain exact solutions. Figure 9 displays the exact solution of Example 4.5 for various values of a (0, 0.2, 0.4, 0.6, 0.8), with y represented on the vertical axis and x on the horizontal axis. The equation for the exact solution is given as $at + xt^2$, where a is a parameter that influences the behavior of the solution. Analyzing the graph in Figure 9, we observe that as the value increases, the corresponding y values also increase. This relationship indicates that higher values of a result in larger magnitudes of y. The curve for $a = 0.8$ exhibits the highest y values, while the curve for $a = 0$ demonstrates the smallest values. This trend suggests that the parameter a significantly impacts the amplitude or magnitude of the solution.

Figure 10 presents a 2D contour plot illustrating the behavior of the exact solution of Example 4.5 for different values of a and x. The vertical axis represents the values of a, while the horizontal axis represents the values of x. The contour lines depicted on the plot represent different levels of the solution, enabling us to visualize how the values change across various regions of the $x - a$ parameter space.

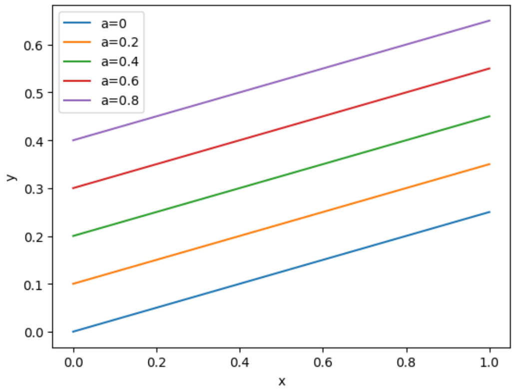


Figure 9. Exact solution of example 4.5 using HPM-LTM method for varying values of t, plotted in Python using Sympy and Matplotlib for different values of a

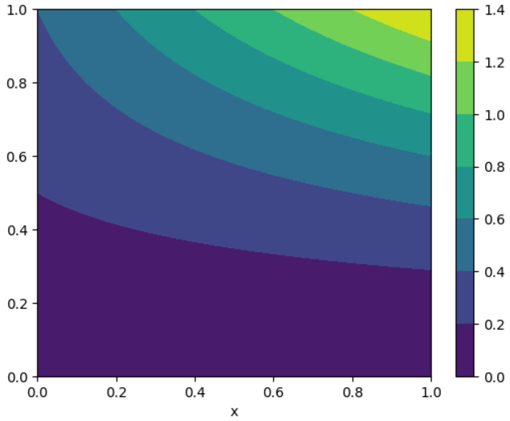


Figure 10. 2D contour plot of the exact solution of example 4.5 using the HPM-LTM method, plotted in Python using Sympy and Matplotlib

By examining the contour lines in Figure 10, we can identify regions where the solution changes more rapidly. In areas where the contour lines are closer together, the solution exhibits significant variations or transitions over a small range of x and a values. Conversely, regions with widely spaced contour lines correspond to areas where the solution changes more gradually or remains relatively constant. This 2D contour plot demonstrates that the behavior of the exact solution is highly dependent on the values of a and x. The contour lines' proximity indicates the solution's sensitivity to changes in these parameters. It suggests that small variations in a and x can substantially change the solution's magnitude or behavior.

Despite the complexity of the solution, the contour plot provides an easily interpretable visual representation of the data. It allows us to identify critical regions, observe trends, and understand the relationship between the solution levels and the values of a and x. This visualization aids in gaining insights into the behavior and properties of the exact solution in the context of Example 4.5.

In our research paper, we compared the efficiency of the HPM-LTM and traditional HPM methods. We performed this comparison by solving example 4.5 using both approaches. While both methods yielded the exact solution, our observations indicated that the HPM-LTM method significantly outperformed the HPM method. Using the HPM-LTM method, we achieved the desired result with just seven blocks of code. In contrast, the HPM method required 48 blocks of code. This substantial difference in coding requirements demonstrates the efficiency gain offered by the HPM-LTM method. The improved efficiency of the HPM-LTM method is attributed to its optimized design, which provides a streamlined and less convoluted approach to problem-solving. The HPM-LTM method simplifies coding by eliminating unnecessary complexities and reducing coding requirements.

Our study unequivocally demonstrates that the HPM-LTM method is superior in efficiently and effectively solving complex problems compared to traditional methods. Our research findings highlight the advantages of adopting the HPM-LTM method for improved efficiency and effectiveness in problem-solving endeavors.

5. CONCLUSION

The proposed HPM-LTM method offers a novel and effective approach for solving linear and nonlinear partial differential equations, as demonstrated through illustrative heat and wave equations examples. By combining the Homotopy Perturbation Method with the Laplace Transform Method, the HPM-LTM method can handle nonlinear terms in equations while avoiding the limitations of traditional methods. Furthermore, Python's computational capabilities enable fast and accurate implementation of the HPM-LTM method. It is a practical computational tool for solving PDE problems in various domains, especially engineering, where efficient calculation time and large data volumes are essential. This study contributes to the advancement of computational tools for solving PDEs, and future research could explore the potential of the HPM-LTM method for more complex PDE problems. In summary, the HPM-LTM method offers a promising approach for solving PDEs with high accuracy and efficiency.

6. REFERENCES

1. He J-H. Homotopy perturbation technique. Computer methods in applied mechanics and engineering. 1999;178(3-4):257-62. https://doi.org/10.1016/S0045-7825(99)00018-3
2. He J-H. Homotopy perturbation method for solving boundary value problems. Physics letters A. 2006;350(1-2):87-8. https://doi.org/10.1016/j.physleta.2005.10.005
3. He J-H. Homotopy perturbation method: a new nonlinear analytical technique. Applied Mathematics and computation. 2003;135(1):73-9. https://doi.org/10.1016/S0096-3003(01)00312-5
4. Shateri A, Moghaddam MM, Jalili B, Khan Y, Jalili P, Ganji DD. Heat transfer analysis of unsteady nanofluid flow between moving parallel plates with magnetic field: Analytical approach. Journal of Central South University. 2023;30(7):2313-23. https://doi.org/10.1007/s11771-023-5388-3
5. LeVeque RJ. Finite difference methods for ordinary and partial differential equations: steady-state and time-dependent problems: SIAM; 2007.
6. Khater M, Anwar S, Tariq KU, Mohamed MS. Some optical soliton solutions to the perturbed nonlinear Schrödinger equation by modified Khater method. AIP Advances. 2021;11(2). https://doi.org/10.1063/5.0038671
7. Sedighi HM, Shirazi KH, Zare J. An analytic solution of transversal oscillation of quintic non-linear beam with homotopy analysis method. International Journal of Non-Linear Mechanics. 2012;47(7):777-84. https://doi.org/10.1016/j.ijnonlinmec.2012.04.008
8. El-Shahed M. Application of He's homotopy perturbation method to Volterra's integro-differential equation. International Journal of Nonlinear Sciences and Numerical Simulation. 2005;6(2):163-8. https://doi.org/10.1515/IJNSNS.2005.6.2.163
9. Van Der Walt S, Colbert SC, Varoquaux G. The NumPy array: a structure for efficient numerical computation. Computing in science & engineering. 2011;13(2):22-30. 10.1109/MCSE.2011.37
10. Nadeem M, Li F. He–Laplace method for nonlinear vibration systems and nonlinear wave equations. Journal of Low Frequency Noise, Vibration and Active Control. 2019;38(3-4):1060-74. https://doi.org/10.1177/1461348418818
11. Filobello-Nino U, Vazquez-Leal H, Sarmiento-Reyes A, Cervantes-Perez J, Perez-Sesma A, Jimenez-Fernandez V, et al. Laplace transform–homotopy perturbation method with arbitrary initial approximation and residual error cancelation. Applied Mathematical Modelling. 2017;41:180-94. https://doi.org/10.1016/j.apm.2016.08.003
12. Liu Z-J, Adamu MY, Suleiman E, He J-H. Hybridization of homotopy perturbation method and Laplace transformation for the partial differential equations. Thermal Science. 2017;21(4):1843-6. https://doi.org/10.2298/TSCI160715078L
13. Jalili B, Ganji AD, Jalili P, Nourazar SS, Ganji D. Thermal analysis of Williamson fluid flow with Lorentz force on the stretching plate. Case Studies in Thermal Engineering. 2022;39:102374. https://doi.org/10.1016/j.csite.2022.102374
14. Jalili B, Sadighi S, Jalili P, Ganji DD. Numerical analysis of MHD nanofluid flow and heat transfer in a circular porous medium containing a Cassini oval under the influence of the Lorentz and buoyancy forces. Heat Transfer. 2022;51(7):6122-38. https://doi.org/10.1002/htj.22582
15. Jalili B, Aghaee N, Jalili P, Ganji DD. Novel usage of the curved rectangular fin on the heat transfer of a double-pipe heat exchanger with a nanofluid. Case Studies in Thermal Engineering. 2022;35:102086. https://doi.org/10.1016/j.csite.2022.102086
16. Endalew MF, Sarkar S. Numerical exploration of forced convection hydromagnetic hyperbolic tangent nanofluid flow over a permeable wedge with melting heat transfer. Scientific Reports. 2023;13(1):3515. https://doi.org/10.1038/s41598-023-30656-2
17. Chamkha A, Ismael M, Kasaeipoor A, Armaghani T. Entropy generation and natural convection of CuO-water nanofluid in C-shaped cavity under magnetic field. Entropy. 2016;18(2):50. https://doi.org/10.3390/e18020050
18. Jalili B, Mousavi A, Jalili P, Shateri A, Domiri Ganji D. Thermal analysis of fluid flow with heat generation for different

logarithmic surfaces. International Journal of Engineering, TRANSACTIONS C: Aspects. 2022;35(12):2291-6. 10.5829/IJE.2022.35.12C.03

19. Jalili P, Narimisa H, Jalili B, Shateri A, Ganji D. A novel analytical approach to micro-polar nanofluid thermal analysis in the presence of thermophoresis, Brownian motion and Hall currents. Soft Computing. 2023;27(2):677-89. https://doi.org/10.1007/s00500-022-07643-2
20. Jalili B, Jalili P, Shateri A, Ganji DD. Rigid plate submerged in a Newtonian fluid and fractional differential equation problems via Caputo fractional derivative. Partial Differential Equations in Applied Mathematics. 2022;6:100452. https://doi.org/10.1016/j.padiff.2022.100452
21. Jalili P, Jalili B, Shateri A, Domiri Ganji D. A novel fractional analytical technique for the time-space fractional equations appearing in oil pollution. International Journal of Engineering, TRANSACTIONS C: Aspects. 2022;35(12):2386-94. 10.5829/IJE.2022.35.12C.15
22. Abbasbandy S. Iterated He’s homotopy perturbation method for quadratic Riccati differential equation. Applied Mathematics and Computation. 2006;175(1):581-9. https://doi.org/10.1016/j.amc.2005.02.014
23. Kiusalaas J. Numerical methods in engineering with Python 3: Cambridge university press; 2013.
24. Johansson R. Numerical Python: Scientific Computing and Data Science Applications with Numpy, SciPy and Matplotlib, Apress, Berkeley, CA, 2019. DOI: https://doi org/101007/978-1-4842-4246-9.
25. Hunter JD. Matplotlib: A 2D graphics environment. Computing in science & engineering. 2007;9(03):90-5. 10.1109/MCSE.2007.55
26. Liao S, Tan Y. A general approach to obtain series solutions of nonlinear differential equations. Studies in Applied Mathematics. 2007;119(4):297-354. https://doi.org/10.1111/j.14679590.2007.00387.x
27. Momani S, Odibat Z. Homotopy perturbation method for nonlinear partial differential equations of fractional order. Physics Letters A. 2007;365(5-6):345-50. https://doi.org/10.1016/j.physleta.2007.01.046
28. Kumar Mishra H, Nagar AK. He-Laplace method for linear and nonlinear partial differential equations. Journal of Applied Mathematics. 2012;2012. https://doi.org/10.1155/2012/180315
29. Shijing D, Hongru C, Xudong W, Deshi W, Yongyong Z. Modal optimization design of supporting structure based on the improved particle swarm algorithm. International Journal of Engineering, TRANSACTIONS A: Basics. 2022;35(4):740-9. 10.5829/IJE.2022.35.04A.14
30. Jalili P, Shateri A, Ganji AM, Jalili B, Ganji DD. Analytical analyzing mixed convection flow of nanofluid in a vertical channel using python approach. Results in Physics. 2023;52:106908. https://doi.org/10.1016/j.rinp.2023.106908
31. 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: Applications. 2022;35(5):1082-8. 10.5829/IJE.2022.35.05B.23
32. Jalili B, Ganji AM, Shateri A, Jalili P, Ganji DD. Thermal analysis of non-Newtonian visco-inelastic fluid MHD flow between rotating disks. Case Studies in Thermal Engineering. 2023;49:103333. https://doi.org/10.1016/j.csite.2023.103333
33. Marchione F. Analytical investigation on the stress distribution in structural elements reinforced with laminates subjected to axial loads. International Journal of Engineering, TRANSACTIONS A: Basics. 2022;35(4):692-7. 0.5829/IJE.2022.35.04A.08
34. Agrawal V, Khairnar H. Analytical Modeling of Heat Transfer Coefficient Analysis in Dimensionless Number of an Electric Parking Brake Using CFD. International Journal of Engineering, TRANSACTIONS B: Applications. 2023;36(2):276-88. 10.5829/IJE.2023.36.02B.08
35. Lap-Arparat P, Tuchinda K. Computational Study of Excitation Controlling Parameters Effect on Uniform Beam Deformation under Vibration. International Journal of Engineering, TRANSACTIONS A: Basics. 2023;36(1):60-70. 10.5829/IJE.2023.36.01A.08
36. Baughman G, Sharma P. Entropic analysis of protein oscillations through langevin equations & fokker-planck equations. Journal of Human, Earth, and Future. 2022;3:56-70. 10.28991/HEF-SP2022-01-05
37. Sutantyo T, Ripai A, Abdullah Z, Hidayat W, Zen FP. Soliton-like solution on the dynamics of modified Peyrard-bishop DNA model in the thermostat as a bio-fluid. Emerg Sci J(ISSN: 2610-9182). 2022;6(4). 10.28991/ESJ-2022-06-04-01
38. Alrwashdeh SS, Ammari H, Madanat MA, Al-Falahat AaM. The effect of heat exchanger design on heat transfer rate and temperature distribution. Emerging Science Journal. 2022;6(1):128-37. 10.28991/ESJ-2022-06-01-010

Persian Abstract

چکیده

این تحقیق یک روش دقیق و نوآورانه به نام روش اختلالی هماتوپی - روش تبدیل لاپلاس (HPM-LTM) که با استفاده از زبان پایتون پیاده‌سازی شده است، برای حل کارآمد معادلات دیفرانسیل جزئی خطی و غیرخطی (PDEs) معرفی می‌کند. با ترکیب تکنیک اختلالی هماتوپی (HPM) با روش تبدیل لاپلاس(LTM) ، روش ما با موفقیت به معاملات مهمی که توسط معادلات دارای اجزای غیرخطی ایجاد می‌شود، پاسخ می‌دهد. با استفاده از چند جمله‌ای‌های هس، رویکرد HPM-LTM با دقت اصطلاحات غیرخطی را مدیریت می‌کند و به راه‌حل‌های دقیق و قابل اعتماد منجر می‌شود. برای نشان دادن کارایی روش ما، ما آن را به طور گسترده در پنج سناریو نماینده ازPDE ، از جمله معادلات حرارت و موج، به کار می‌بریم. نتایج جامع ما دقت و قابلیت اعتماد قابل توجه روش HPM-LTM را تأیید می‌کند و برتری آن نسبت به روش‌های معمولی که نیاز به فرضیات محدود کننده یا گسسته‌سازی دارند و ممکن است خطاهای گردایی ایجاد کنند، را مورد تأکید قرار می‌دهد. علاوه بر این، روش ما محدودیت‌های ناشی از خطاهای عددی موجود در تکنیک‌های هماتوپی سنتی را برطرف می‌کند. قدرت، کارایی و انعطاف‌پذیری روش پیشنهادی ما از طریق استفاده موفق از آن در مجموعه‌ای گسترده از مسائل PDE در زمینه‌های مختلف تأیید می‌شود. این تحقیق یک مشارکت مهم به توسعه ابزار محاسباتی قدرتمند برای حل مسائل متنوع PDE با تأکید خاص بر رشته مهندسی معرفی می‌کند.

IJE TRANSACTIONS B: Applications Vol. 37 No. 02, (February 2024) 365-376

International Journal of Engineering

Journal Homepage: www.ije.ir

Effect of SiO_2 Concentration and Time on Stability of TiO_2 Zeolite Nanocomposite Membrane in Light Gas Dehumidification

G. R. Fouladi[a], H. Mazaheri*[a], A. Marjani[b]

[a] *Department of Chemical Engineering, Arak Branch, Islamic Azad University, Arak, Iran*
[b] *Department of Chemistry, Arak Branch, Islamic Azad University, Arak, Iran*

PAPER INFO

Paper history:
Received 20 August 2023
Received in revised form 19 September 2023
Accepted 24 September 2023

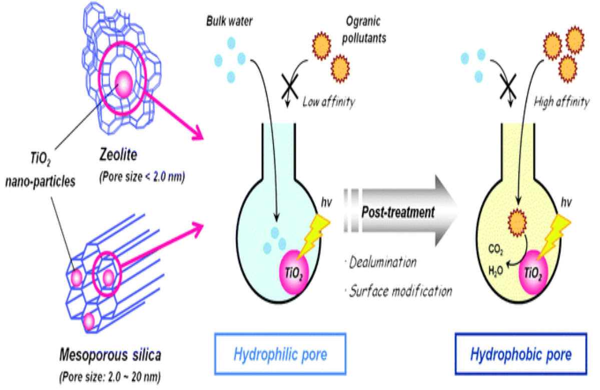
Keywords:
Dehumidification
SiO_2 Concentration
Nanocomposite Membrane
TiO_2 Zeolite
Stability

ABSTRACT

The main goal of this research is to use zeolite TiO_2 nanocomposite membranes in order to remove moisture from gas. For this reason, a certain TiZ-V membrane was selected and manufactured as a standard membrane, which was the result of the initial assessment of a suitable membrane for gas dehumidification, and this membrane was used as a standard to measure the effect of manufacturing parameters. The findings showed that increasing the concentration of SiO_2 had greatest effect on increasing the water flux of the membrane due to the effect of increasing the reaction time of the vapor phase carrier and reducing the selectivity drop at higher pressures. Also, the experiments of changing the relative humidity of the feed have shown the improving efficiency of the membrane in relative humidities lower than 80%, so that under the conditions of lower relative humidity, the selectivity of the membrane has increased. Another positive point found is a slight change in the selectivity efficiency of the membrane with respect to different relative humidities. This case showed the stability of membrane performance under the different conditions of humidity of the feed gas. Next, in order to increase the performance of membrane as much as possible, the sweeper gas was added from the inside of membrane. Increase in the sweeper gas, which increases the water concentration gradient and decreases the gas concentration gradient on the sides of membrane wall, increases the selectivity of membrane to the highest level of 543.

doi: *10.5829/ije.2024.37.02b.12*

Graphical Abstract



NOMENCLATURE		Greek Symbols	
P	bar	Sw	mL/s
V	mL	$RH_R\%$	relative humidity of residual
T	°C	$RH_F\%$	relative humidity of feed
Q	L/h	GPU	permeability
a	selectivity	P	$mol/m^2.s.pa$
P_{avg}	pa	ρ	Density (kg/m^3)
g	Gravity(m/s^2)	τ	Lattice relaxation time
n	mol	Y^G	mole fraction of gas
Y^w	mole fraction of water	t	time (min)

---

*Corresponding Author Email: *H.Mazaheri@iau.arak.ac.ir*
(H. Mazaheri)

Please cite this article as: Fouladi GR, Mazaheri H, Marjani A. Effect of SiO_2 Concentration and Time on Stability of TiO_2 Zeolite Nanocomposite Membrane in Light Gas Dehumidification, International Journal of Engineering, Transactions B: Applications. 2024;37(02):365-76.

1. INTRODUCTION

Because natural gas is one of the sources of energy production in the transmission industry and domestic use, its production and distribution is very important (1). Natural gas is mainly composed from methane (80 to 90 percent), a small amount of ethane-propane and butane. In addition, gas compounds often include other compounds such as H_2O and CO_2 that the presence of these compounds in gas causes problems during transportation and consumption (2-4). Because natural gas often leaves the reservoirs as saturated from water, A slight change in gas pressure or temperature causes water condensation or formation of hydrates (5). Hydrates are compounds very similar to ice which are obtained from the bonds between water and ice molecules (6). The formation of hydrates during gas transmission causes blockage of the pipeline and consequently, permanent interruption of gas flows (7). In addition to hydrate formation, the presence of liquid water in pipelines can cause pipe corrosion; therefore, it is necessary to be removed from the composition of natural gas (8). Removal of moisture from natural gas is due to the need to prevent from reducing the calorific value of gas and purifying gas in order to achieve the necessary standards for delivering gas to Petrochemical industries and household uses (9, 10). Also, in the facilities department in the gas industry, the presence of water in gas pipelines can cause two main problems; the presence of small amounts of water along with the presence of some sour gases causes widespread acid corrosion in gas transmission pipelines. This issue would cause the imposition of costs and the possibility of risks due to gas leakage (11, 12).

Also, water is one of the main bases for the formation of gas hydrates (13), and formation of solid methane hydrates in pipes containing gas in different parts of the facility can change the flow pattern, pressure changes and also, with its rapid growth, it can stop the gas transmission operation due to the blockage of the pipe and the possibility of its bursting (14). As a result, dehumidification of natural gas from the formation of gas hydrates helps proper transmission and reduction of corrosion. There are various methods for dehumidifying natural gas on an industrial scale (15, 16), the most important of which are as follows:

1. Direct cooling
2. Surface adsorption
3. Absorption by liquid

Dehumidification of natural gas is done in two main approaches: Towers filled with surfactants and the wider method that is ethylene glycol dehumidification units (17).

The first method includes towers filled with solid surface absorbents such as silica gel or alumina (18), which have the problems as follows:

The need for several towers (on average 3 or 4 towers); restriction of use due to weight and large dimensions, the need for permanent and accurate operation controls, extensive destruction and high price of absorbent, also this method often has the problems of ethylene glycol dehumidification units (19).

A more widespread method of dehumidification unit is tri-ethylene glycol (TEG) (20). This method only involves the use of two towers that is contacted with ethylene glycol which is a liquid of water absorbent in the gas absorption tower in the form of counter flow on the bubble cap trays and then in the solvent recycling operations disposal tower, is required to separate moisture from tri-ethylene glycol (21). Some of the main problems with glycolic dehumidifiers are:

1. Absorbent wide flammability; Result: safety issues.
2. Complex unit operation in absorption and disposal towers; Result: high investment, operating and maintenance costs.
3. Energy supply of towers and facilities; Result: consuming part of gas as fuel.
4. Requiring solvent warehouse and storage, replacing, neutralizing and reproducing it; Result: the need for a large space.
5. Emission of hazardous volatile organic compounds (VOCs), especially benzene, toluene, ethylbenzene, and xylenes (BTEX); Result: release of 40 tons of VOCs per year from a typical dehumidification unit that half of which is BTEX (22-24).

If the goal is to dehumidify of large volume of gas to reach a dew point of 40 to 140 degrees Fahrenheit, liquid absorption systems are more economical (25). But if the dew point is above 180 degrees Fahrenheit, membrane absorption units are more suitable (26). Due to the disadvantages of moisture, extensive studies are carried out in the field of investigating the effect of various parameters such as the type of adsorbent, feed flow intensity, temperature and process pressure on the moisture absorption performance of the gas flow (27). Therefore, temperature, pressure, moisture concentration in feed, Intensity of feed flow (independent variable) as variables affecting moisture absorption performance in a nanostructure membrane system were investigated (23). In terms of the growing trend of problems caused by the method of dehumidification conventionally, the necessity of conducting research by promising technologies such as Nano is more apparent.

Among the various techniques, membrane filtration is a more effective and efficient one for separating water from gas, especially oil emulsions. The advantages of membrane separation are high selectivity, low maintenance cost, low required space, high efficiency, and no need for chemical additives (28).

Among the most important advantages of Nano composites are better mechanical properties compared to normal composites (reinforcement in Nano composites is more bidirectional than unidirectional), increasing thermal stability and self-extinguishing, preventing the penetration of gases such as oxygen, steam He mentioned water, helium, carbon dioxide and organic vapors such as ethyl acetate (due to the formation of a porous and tortuous path due to the presence of clay) (29).

Also, nanocomposites are lighter than normal composites, and this feature makes them competitive with other materials with special applications (29).

This study was conducted regarding the possibility of using TiO_2 zeolite nanocomposite membranes to remove moisture from gas with natural gas and nitrogen approach and about gas membrane dehumidification. In this way, the present work has been completely innovative in terms of the selected method and the type of surface material used in the fabrication of TiZ-V membrane on NAA zeolite hollow fibers.

2. MATERIALS AND METHODS

2. 1. Synthesis of Preparation of Titanium Dioxide Nanowire by Eutectic Method

Due to the methods of preparation of nanostructures and also checking the condition of the equipment and available facilities, eutectic method has been used for fabrication of titanium nano oxide. Because the selected equipment used in this method is relatively cheap, the need for low temperature in the process, low energy consumption and compatible with the environment. To prepare titanium dioxide One-dimensional nano structures (nanowire) by eutectic method, a mixture of disodium phosphate salts (Na_2HPO_4), sodium chloride (NaCl) and titanium dioxide raw material (TiO_2) with specific weight ratios (1:4:1) was prepared and ground in a mortar, then transferred to the furnace in the porcelain crucible and the reaction was performed at temperature 825 °C for 8 hours. After cooling, the crucible is removed from the furnace and then several times washing was done with boiled distilled water. During this step, the excess salts in the products were washed. At finally, the sediments from washing were dried in the furnace (Figure 1).

Figure 1. Different stages of synthesis by Eutectic method

2. 2. Synthesis of Zeolite-NaA Membrane

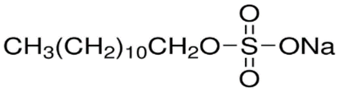
The synthesis of zeolite was done by micro emulsion method and according to the report of Shams et al. (30). The surfactant used in this work was sodium di-decyl sulfate (SDS)$_5$ with the general formula $CH_3(CH_2)_{11}OSO_3Na$ which is an anionic surfactant with HLB=40 and its chemical structure is shown in Figure 2.

The numerical value of HLB plays an important role in deciding to choose the type of surfactant and indicates the polarity and non-polarity ratio of surfactant groups. To form a water-in-oil emulsion, HLB is needed between 5.3 to 8. But, SDS has a general mode and is used for the most tests. In order to prepare micellar solution, surfactant (SDS) and auxiliary surfactant (butanol) with a weight ratio of 2:1 are mixed with solvent (Hexane). Auxiliary surfactant is used to neutralize electrostatic charges caused by side co-positioning of the charged heads of the ionic surfactant and stabilizing the micelles. To dissolve SDS in hexane, the resulting mixture is kept night and day in a closed glass container and mixed by a magnetic stirrer to achieve complete uniformity. Then the composition of zeolite is prepared with the molar ratio of Al_3O_3, $2.1SiO_2$, $3Na_2O$, $150H_2O$. In short, 2g of aluminum sulfate and 0.58g of sodium hydroxide is dissolved in 6.22g of deionized water and will be poured into syringe A. 1.41g of sodium silicate solution is poured into syringe B.

The contents of two syringes were added drop by drop to the micellar solution under ultra sonication. The weight ratio of mixing materials was 3H:1S:1Z, where H, S and Z represent solvent, the mixture of surfactants and aluminosilicate gel (the total weight of aluminum solution and silicate solution) respectively. Mixing of the ingredients continued for 30 minutes and then the resulting mixture was transferred to a steel reactor which its inner body was covered with Teflon and placed in an oven at a temperature of 75°C for 6 hours. After this period, the contents of the reactor were removed and centrifuged to separate zeolite particles from the water and oil mixture, which are two separated phases. Then zeolite particles were washed several times with water and acetone to reduce the pH of the water after washing. Then zeolite particles were placed inside the oven to dry at a temperature of 120°C for 12 hours. In order to calcine and remove the remaining surfactants, zeolite particles are placed in furnace at a temperature of 550°C for 5 hours.

2. 3. SiO_2 Deposition on TiO_2

In this step, 1g of particles (TiO_2) obtained from step 1 is dispersed in deionized water by a homogenizer and the resulting



$$CH_3(CH_2)_{10}CH_2O-\overset{\overset{\displaystyle O}{\|}}{\underset{\underset{\displaystyle O}{\|}}{S}}-ONa$$

Figure 2. The chemical structure of sodium di-decyl sulfate

mixture is placed in a water bath at a temperature of 90°C. Then the amount of 3.7 g of sodium silicate (the weight ratio of $\frac{SiO_2}{TiO_2}=\frac{1}{1}$) is added to the solution. At this time, 0.1M hydrochloric acid solution is added drop by drop to the solution until the pH of the solution reaches 0.5. Stirring of the solution continues for 1h and after this period, the solution is cooled at room temperature and the covered particles (TiO_2) with SiO_2 are separated with a sieve. (TiO_2) @SiO_2 particles were washed and dried at 50°C. These particles were called P−TiS.

2. 4. P−TiS Covering with NaA Zeolite

In order to coat P-TiS particles by zeolite, the vapor phase transport method (VPT) will be used, which is a method in synthesis of high silica zeolites. Matsukata et al. (31) reported that dry alum inosilicate gel was converted to MFI when exposed to water vapors and volatile amines. In fact, in this method, amorphous alum inosilicate gel is crystallized in contact with vapors of a structure-forming material without direct contact with it. In this research, this method was used to integrate the SiO_2 layer in zeolite. In short, the synthesized zeolite particles in step 2 were mixed with 30mL of toluene and placed in a closed container on a magnetic stirrer for one day and night. Then P-TiS particles were added to the mixture under vigorous stirring by a homogenizer. The weight ratio of P−TiS to zeolite is 1:1. Stirring was continued for 10 minutes. Then the mixture was placed statically to deposit TiO_2 particles covered with zeolite. The resulting particles were placed under vacuum and at ambient temperature for 2 hours to evaporate the remaining toluene. The resulting solid was transferred to a steel reactor with Teflon inner body. An overview of the used reactor can be seen in Figure 3. The particles were placed on the cylinder inside the reactor and a solution containing 4g of deionized water and 2g of trimethylamine was poured around it. Reactor was kept at a temperature of 140°C for 3 days. Then the resulting particles were removed from the reactor and dried at 40°C. These particles are named with TiZ−V (32).

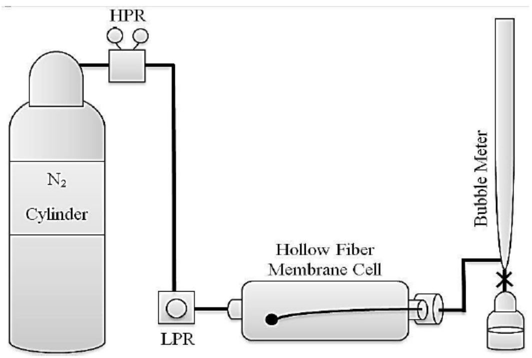


Figure 3. Design of gas test device of the hollow fiber membrane

The entering gas from the cylinder enters to the chamber with a certain pressure. This gas goes from the side of the shell by passing through membrane wall to the inside of the hollow fiber and from there is directed to the bubble flow meter. According to the amount of permeated gas in the bubble flowmeter, gas permeability will be calculated (see Figure 4).

In order to determine the selectivity of membrane, the data obtained from the experiments are used as input for the calculations include; different pressures of the inlet gas (P[bar]), a certain volume of the bubble flow meter (V[mL]), The gas flow time in the bubble flow meter for a given volume (t[s]), ambient temperature and vapor generator tank temperature (T [°C]), relative humidity of feed RH_F, relative humidity of penetrant flow RH_P, relative humidity of residual RH_R and feed flow (L/h) which are used to measure the dehumidification performance of the constructed membrane. It should be noted that the considered standard operational conditions in performance tests include relative humidity of feed, 80%, the gas flow is 50 [L/h] and without sweep flow which, these conditions in all operations except those mentioned in the measurement section of the effect of change operating parameters are followed by default.

3. RESULTS AND DISCUSSIONS

3. 1. SiO_2 Concentration

The effect of change the concentration of SiO_2 on the results of gas, water and performance tests are shown in Figures 5, 6 and 7, respectively.

Regarding the $1SiO_2$ membrane, the water permeability of the membrane has decreased with a decrease in concentration of SiO_2, which is also

Figure 4. Image of the constructed gas dehumidification membrane device

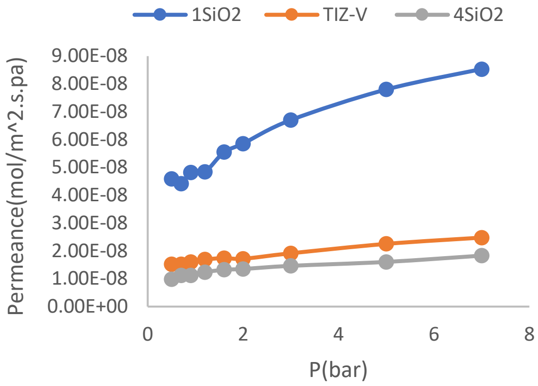


Figure 5. Diagram of the effect of change the SiO_2 concentration on gas test results

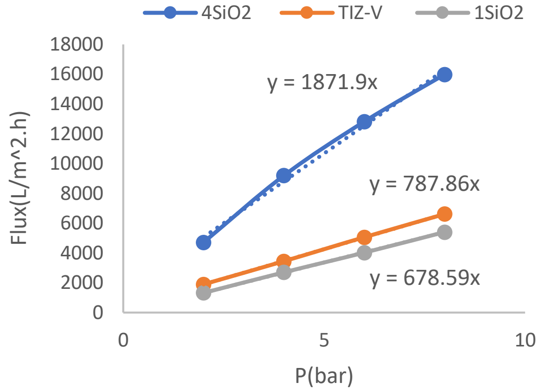


Figure 6. Diagram of the effect of change the SiO_2 concentration on the results of water test

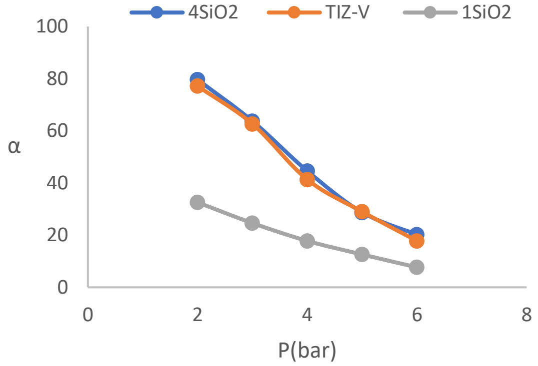


Figure 7. Diagram of the effect of change the SiO_2 concentration on performance test results

considered an unfavorable situation. The decrease in concentration of the silicate phase in the SiO_2 aqueous solution has caused a decrease in the hydrophilicity of the membrane and, as a result, a decrease in the water permeability of the membrane. The results of $1SiO_2$ membrane efficiency test are presented in Table 1.

TABLE 1. Permeability results of zeolite membrane with variations of $1SiO_2$ concentration

P(bar)	$P_{avg}(pa)$	N (mol)	$P(mol/m^2.s.pa)$
0/1	106325	-06E2/2442	-06E1/766
0/3	116325	-06E6/7191	-06E1/762
0/5	126325	-05E1/1174	-06E1/758
0/7	136325	-05E1/56294	-06E1/757
0/9	146325	-05E2/00279	-06E1/751
1/2	161325	-05E2/66429	-06E1/747
1/6	181325	-05E3/54392	-06E1/743
2	201325	-05E4/4189	-06E1/738
3	251325	-05E6/61227	-06E1/734
5	351325	0/000109702	-06E1/726
7	451325	0/000153287	-06E1/723

Regarding the $4SiO_2$ membrane, it shows an increase in water permeability of the membrane with the increase of SiO_2 concentration. Increasing the concentration of the silicate phase in the mixture of SiO_2 has increased the hydrophilicity of the membrane and, as a result, increased the water permeability of the membrane. The results of the performance test of $4SiO_2$ membrane are presented in Table 2.

3. 1. 1. The Effect of Decrease in the SiO_2 Concentration

According to the comparison between the concentrations of SiO_2 in (Figure 7), respect to TiZ-V membrane, it is concluded that the decrease in the concentration of SiO_2 increased the gas permeability. An increase in gas permeability is considered as an unfavorable state because membrane must have minimum gas permeability and maximum water permeability. The cause of this phenomenon is due to the decrease in the concentration of SiO_2 and as a result, the failure to complete the cross-linking process on the

TABLE 2. Permeability results of zeolite membrane with variations of $4SiO_2$ concentration

$P_{avg}(pa)$	n(mol)	$P(mol/m^2.s.pa)$
126325	-07E2/08963	-08E4/078
136325	-07E2/90998	-08E4/056
146325	-07E3/78164	-08E4/100
161325	-07E5/11107	-08E4/156
181325	-07E6/86339	-08E4/185
201325	-07E8/60661	-08E4/199
251325	-06E1/33895	-08E4/355
351325	-06E2/33977	-08E4/566
451325	-06E3/39365	-08E4/730

surface of membrane, which causes the cross-linking layer of membrane, especially at higher pressures (according to Figure 5) loses the possibility of resistance to prevent penetration of nitrogen molecules and then, the permeability of nitrogen gas increases.

Perusing the diagram in (Figure 6), regarding membrane $1SiO_2$, shows that the decrease in water permeability of membrane have been with the decrease in the concentration of SiO_2. This is also considered as an unfavorable state. The decrease in the concentration of the silicate phase in the aqueous solution of SiO_2 has caused a decrease in the hydrophilicity of membrane and consequently, a decrease in the water permeability of membrane. Membrane performance test results of $1SiO_2$ is presented in Table 3.

According to Table 3, regarding the state of penetrant, it indicates that water component is lower compared to gas even in the lowest measured pressure. This state means an increase in gas permeability from the penetrant side, which means an increase in gas loss from the penetrant side in the form of wet gas which, corresponds to the result of gas test of this membrane. Also, the lower water content in the penetrant corresponds to the result of water test of this membrane, i.e., the decrease in water permeability. These two unfavorable states have caused a general decrease in the selectivity of the mixed gas of membrane. Relative humidity change diagram of membrane $1SiO_2$ is also presented in Figure 8. Finally, Figure 9 presents the gas and water permeability of the $1SiO_2$ membrane at different pressures and their corresponding selectivity.

3. 1. 2. The Effect of Increase of the SiO_2 Concentration ($4SiO_2$ Membrane)

According to the comparison between the concentrations of SiO_2 in (Figure 7), relative to TiZ-V membrane, it is concluded that the increase in the concentration of SiO_2 has caused a slight decrease in gas permeability. The increase in SiO_2 provides the required values for the reaction with the existing value of $P{-}TiS$, but its higher values will not cause more crosslinking of the network due to the no reaction with $P{-}TiS$; therefore, as the results show, the major no change in permeability values of the $4SiO_2$ membrane was due to the major no change in the crosslinking of the selective layer compared to TiZ-V membrane.

TABLE 3. Membrane performance test results of $1SiO_2$

p(bar)	n_F	n_p	Y^w	Y^G	a
2	5.797E-04	1.120E-05	0.460	0.540	8.32
3	5.827E-04	1.41973E-05	0.395	0.605	0.25
4	5.877E-04	1.92646E-05	0.320	0.680	1.18
5	5.981E-04	1.96823E-05	0.251	0.749	8.12
6	6.197E-04	5.12518E-05	0.169	0.831	7.7

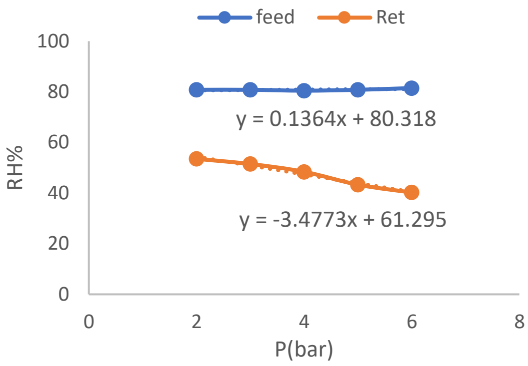


Figure 8. Diagram of changes in relative humidity of residue versus relative humidity of feed in the $1SiO_2$ membrane

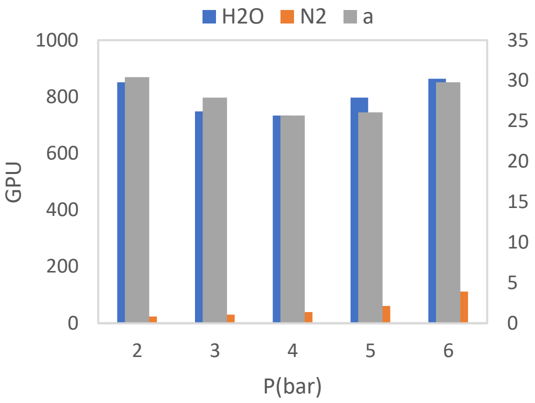


Figure 9. Diagram of gas and water permeability values of $1SiO_2$ membrane and its corresponding selectivity

The comparison of gas permeability values in $1SiO_2$ membrane compared to $4SIO_2$ was done and data are summarized in Table 4.

Perusing the diagram in Figure 6, regarding to the $4SiO_2$ membrane, shows an increase in membrane water

TABLE 4. Comparison of gas permeability values of 1SiO2 membrane compared to 4SiO2

$P_{avg}(pa)$	$1SiO_2, P(mol/m^2.s.pa)$	$4SiO_2, P(mol/m^2.s.pa)$	$1SiO_2/4SiO_2, P(mol/m^2.s.pa)$
126325	-06E1/758	-08E4/078	+01E4/31
136325	-06E1/757	-08E4/056	+01E4/33
146325	-06E1/751	-08E4/100	+01E4/27
161325	-06E1/747	-08E4/156	+01E4/20
181325	-06E1/743	-08E4/185	+01E4/16
201325	-06E1/738	-08E4/199	+01E4/14
251325	-06E1/734	-08E4/355	+01E3/98
351325	-06E1/726	-08E4/566	+01E3/78
451325	-06E1/723	-08E4/730	+01E3/64

permeability with increasing SiO_2 concentration. Increasing the concentration of the silicate phase in SiO_2 mixture has increased the hydrophilicity of membrane and as a result then increased the permeability of water through membrane. The test results of 4SiO_2 membrane performance are presented in Table 5.

Examining Table 5, regarding to the state of penetrant, it shows that water component is higher than the gas from low pressures up to 4 bar pressure, which compared to 1SiO_2 membrane, it is concluded that an increase in the concentration of SiO_2 has improved the state of permeability by increasing its water content. According to the results of relative humidity change diagram of 4SiO_2 membrane presented in Figure 10, the same improvement is clear compared to membrane 1SiO_2. But according to the diagram in Figure 7, regarding to the comparison of performance with standard TiZ-V membrane, a slight improvement has been achieved in selectivity which is similar to relative humidity change graph data in Figure 10. This means that increasing the concentration of SiO_2 increases the water permeability in water test. Finally, Figure 11 presents the gas and water permeability of the 4SiO_2 membrane at different pressures and their corresponding selectivity.

3. 2. The Effect of Operating Conditions

After measuring the effect of different parameters in membrane construction, in this section, this membrane is measured under changes in some operating conditions. For this purpose, the 0.4 TiZ-V membrane has been subjected to the measurement of relative humidity change of feed, the change of the flow rate of dry gas, the effect of the sweeper gas and the test of increasing the operation time.

TABLE 5. 4SiO_2 membrane performance test results

P (bar)	n_F	n_p	Y^w	Y^G	a
2	5.845E-05	1.601E-05	0.674	0.326	3.79
3	5.863E-05	1.78387E-05	0.628	0.372	6.63
4	5.893E-05	2.08073E-05	0.536	0.464	2.44
5	5.952E-05	2.6619E-05	0.427	0.573	6.2 8
6	6.032E-05	3.46833E-05	0.345	0.655	0.20

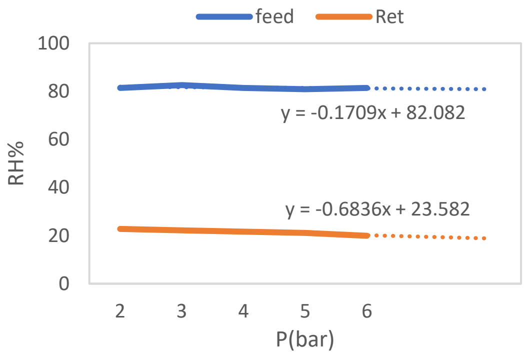


Figure 10. Diagram of changes in the residue relative humidity relative to feed in the 4MPD membrane

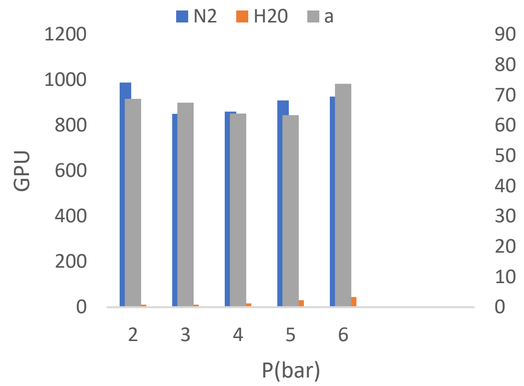


Figure 11. Diagram of gas and water permeability values of 4SiO_2 membrane and its corresponding selectivity

3. 2. 1. Change Relative Humidity of Feed

In this experiment, by change the element temperature in the steam generator, it changes relative humidity of the inlet wet gas to the module. The results of performance tests in relative humidity values of about 50%, 65% and 80% are presented in Table 6. It should be noted that relative humidity data were 100% in this experiment as well as all experiments in the present study which indicates the high humidity of the penetrant flow in TiZ-V membranes and it has been due to the high absorption of water from the humid gas.

According to Table 6, it can be seen that at any given relative humidity of feed, selectivity has decreased with increase of pressure which, is due to the decrease of water component relative to the gas with increase of pressure. It is also seen that decrease of relative humidity of feed from 80% to 50% increased the selectivity value. This is due to the fact that the incoming gas is drier which means less moisture content with the gas.

As it can be seen in Figure 12, the selectivity of membrane has increased with decrease in relative humidity of feed and the remarkable case is small change of the selectivity values with the change of relative humidity of feed. This matter shows the stability of membrane function under the operating conditions of change relative humidity, so that with a decrease in relative humidity, we see an increase in selectivity of membrane and with an increase in relative humidity, decrease in selectivity is slightly changes. This matter can also be seen in Figure 13, which shows the change in relative humidity of the residue.

3. 3. Change the Gas Flow Rate

In this section, by adjusting the flow meter placed on the residue side,

TABLE 6. Test results of relative humidity change of feed

a	Y^G	Y^w	n_p	n_F	RH_R%	$P(bar)$	RH_F%
9.376	0.144	0.856	8.944E-06	5.774E-04	5.7	2	8.49
3.303	0.170	0.830	8.47891E-06	5.779E-04	2.7	4	6.50
0.225	0.218	0782	0.0867E-05	5.786E-04	8.6	6	3.50
5.371	0.144	0.886	1.157E-05	5.800E-04	9.0	2	3.65
2.295	0.140	0.860	1.21519E-05	5.806E-04	9.7	4	3.65
4.211	0.185	0.815	1.29161E-05	5.814E-04	4.7	6	1.65
3.367	0.096	0.904	1.411E-05	5.826E-04	4.10	2	2.80
4.288	0.118	0.882	1.48941E-05	5.834E-04	3.9	4	1.81
2.195	0.165	0.835	1.59014E-05	5.844E-04	2.8	6	6.80

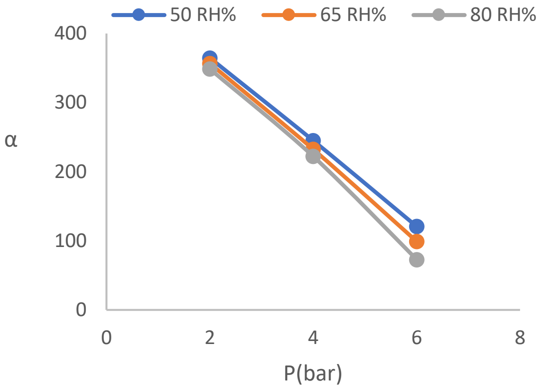


Figure 12. Diagram of selectivity with change of relative humidity of feed

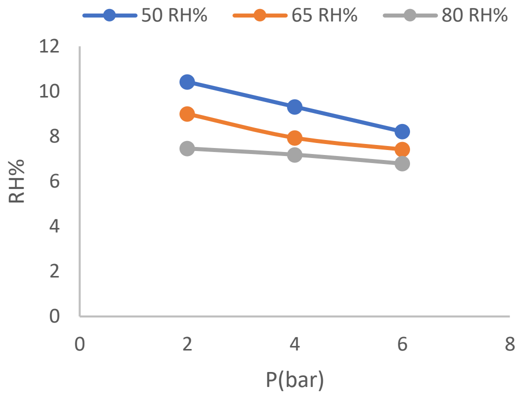


Figure 13. Relative diagram of the residue with change of relative humidity of feed

we set and check the total flow rate of the gas flow in different values. It should be noted that the gas flow rate was 50 [L/h.] in other tests except for this test which is related to the change of gas flow rate. In addition to the flow rate of 50[L/h], values of 100 and 150 have also been investigated. Performance test results for the gas flow rate change test are presented in Table 7.

Relative humidity of feed in this experiment is set at 65%.

The data in Table 7 show the decrease in membrane selectivity with increasing gas flow rate. This matter which has been occurred due to the reduction of the molecules stopping time in the vicinity of membrane in

TABLE 7. Results of the gas flow rate change test

a	Y^G	Y^w	n_p	n_F	RH_R%	$P(\ bar)$	QF
5.371	0.114	0.886	1.157E05	5.800E-04	9.0	2	50
2.295	0.140	0.860	1.21519E05	5.806E-04	9.7	4	50
4.211	0.185	0.815	1.29161E05	5.814E-04	4.7	6	50
4.336	0.125	0.875	2.230E05	1.159E-03	6.11	2	100
2.255	0.158	0.842	2.38157E05	1.161E-03	2.10	4	100
8.188	0.203	0.797	2.56262E05	1.163E-03	3.9	6	100
8.304	0.136	0.864	3.219E05	1.738E-03	2.14	2	150
8.231	0.172	0.828	3.4146E05	1.740E-03	3.13	4	150
1.170	0.220	0.780	3.74936E05	1.743E03	9.11	6	150

order to separate the moisture from the gas is shown in Figure 14. Also, Table 7 shows maintaining the superiority of the difference between water component and the gas in all the pressures and all the investigated flow rates, which is a completely favorable state. Another result obtained from the gas flow rate change test is the possibility of membrane performance with a selectivity of 170 to 304 at investigated high flow rate.

3. 4. The Effect of Sweeper Gas Test

In this section, the effect of adding sweeper gas to the inner side of the hollow fibers is investigated on increasing the selectivity. For this purpose, dry nitrogen is inserted with pressure close to zero at different flow rates from the end side of the module that we blocked by placing the cap in the previous tests. Also, regarding the other operating conditions of this experiment, it should be mentioned that the flow rate and the humidity level of feed are set at 50 [L/h] and 80% respectively. The results of the sweep gas effect test are presented in Table 8.

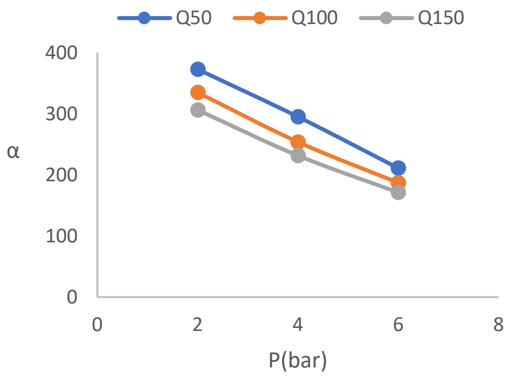


Figure 14. Membrane selectivity graph with change gas flow rate

As can be seen in Table 8, the entry and increase of sweeper gas has a significant effect on increase of membrane selectivity. This phenomenon is due to the role of the sweeper gas in reducing the gas concentration gradient on the outer and inner sides of membrane, so that due to the presence of gas inside the hollow fibers, the external gas has less possibility and tendency to penetrates into membrane. As can be seen in the table about, gas component in penetrant has decreased with the increase in the sweeper gas flow rate, which indicates to this matter. Another positive role of the sweeping gas is in moving the resulting moisture inside the fibers which causes an increase of the water gradient pressure inside and outside of membrane by faster movement and exit of the amount of water inside the fiber. And as a result, the water in the wet flow has a greater possibility and tendency to penetrate into membrane which is a very favorable role. The factor that limits the use of high amounts of sweep gas in industries is the economic debate about the cost of dry gas especially the use of the same type of under operation gas.

Figure 15 shows the changes in relative humidity of the residue relative to feed for this case which shows the decrease of relative humidity of the wet gas from 80% to an average of about 7% during only one membrane phase.

3. 5. Increase the Operation Time Test

In this section, the changes in the performance of membrane selectivity are investigated during the increase in the duration of membrane operations in such a way that operational data has recorded every 30 minutes during 7 hours of continuous operations to obtain the graph of selectivity changes during this period. The operational conditions include feed relative humidity equal to 80%, feed gas pressure equal to 4 bar, gas flow rate equal to 50 [$L/h.$] and without using sweeper gas.

TABLE 8. The results of the sweeper gas effect test

a	Y^G	Y^w	n_p	n_F	$RH_R\%$	p(bar)	$Sw(mL/s)$
3.367	0.096	0.904	1.411E-05	5.826E-04	4.10	2	0
4.288	0.118	0.882	1.448941E-05	5.834E-04	3.9	4	0
2.195	0.165	0.835	1.59014E-05	5.844E-04	2.8	6	0
1.416	0.086	0.914	1.421E-05	5.827E-04	9.8	2	1
5.323	0.107	0.893	1.47108E-05	5.832E-04	5.8	4	1
6.216	0.152	0.848	1.55579E-05	5.840E-04	2.8	6	1
8.494	0.073	0.927	1.409E-05	5.826E-04	6.8	2	5
7.381	0.093	0.907	1.44097E-05	5.829E-04	3.8	4	5
9.276	0.123	0.877	1.50901E-05	5.836E-04	9.7	6	5
1.543	0.067	0.933	1.413E-05	5.826E-04	1.8	2	10
5.418	0.085	0.915	1.44652E-05	5.829E-04	7.7	4	10
1.315	0.110	0.890	1.49611E-05	5.834E-04	4.7	6	10

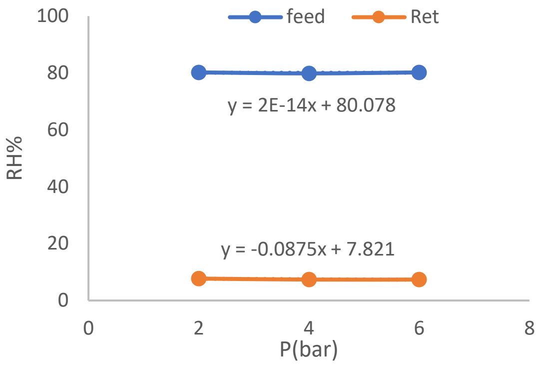


Figure 15. Diagram of changes in relative humidity of the residue relative to feed in Sw10 mode

The results of the increase the operations time test and its changes are presented in Table 9 and Figure 16, respectively. These results indicate the stability of membrane performance over time. As can be seen, this stability of the results is also seen in the removal of moisture from the gas which in all results, the gas moisture has decreased from about 80% to about 90% during only one membrane phase and also regarding the superiority of water component in the penetrant relative to the gas component, there is this stability which indicates the performance of membrane to operate at higher pressures. In addition, it can be seen that the rate of selectivity with an average of 288 in the investigated time period according to the above graph has also had a relatively increasing trend. In particular, as shown in Figure 16, the changes in selectivity of membrane in all investigated times has been varied only between 287.5 and 289.5 which shows the stability of membrane performance with increasing time of operations.

TABLE 9. Test results of increase the operations time

α	Y^G	Y^w	n_p	n_F	$RH_R\%$	t
6.287	0.119	0.881	1.476E-05	5.831E-04	5.9	30
9.288	0.118	0.882	1.47808E-05	5.832E-04	1.9	60
8.287	0.119	0.881	1.47096E-05	5.832E-04	2.9	90
0.288	0.119	0.881	1.46857E-05	5.832E-04	4.9	120
6.288	0.119	0.881	1.46873E-05	5.832E-04	1.9	150
6.287	0.119	0.881	1.47496E-05	5.832E-04	1.9	180
0.288	0.118	0.882	1.47595E-05	5.832E-04	2.9	210
4.288	0.119	0.881	1.47439E-05	5.832E-04	2.9	240
6.288	0.119	0.881	1.48048E-05	5.833E-04	9.8	270
8.287	0.119	0.881	1.46909E-05	5.832E-04	2.9	300
3.289	0.119	0.881	1.47404E-05	5.832E-04	1.9	330
5.287	0.119	0.881	1.467E-05	5.831E-04	4.9	360
3.279	0.118	0.882	1.47591E-05	5.832E-04	1.9	390
0.288	0.119	0.881	1.47657E-05	5.832E-04	2.9	420

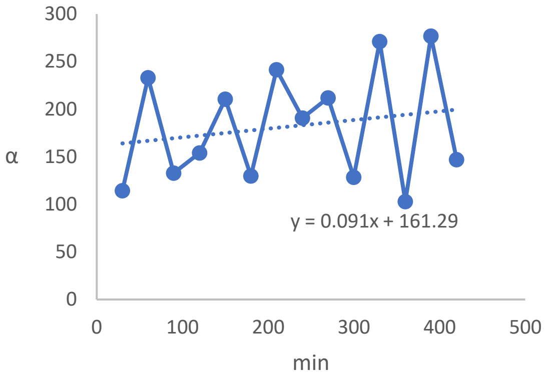


Figure 16. Diagram of selectivity changes with increase the operations time

4. CONCLUSION

The study was conducted on the possibility of using TiO_2 zeolite nanocomposite membranes in order to remove moisture from gas with the approach of natural gas and nitrogen and with the aim of influencing membrane manufacturing parameters.

According to the comparison between the SiO_2 concentrations, compared to the TiZ-V membrane, it is concluded that the decrease in SiO_2 concentration has increased the gas permeability. An increase in gas permeability is considered an undesirable situation because the membrane must have minimum gas permeability and maximum water permeability. The cause of this phenomenon is due to the decrease in the concentration of SiO_2 and as a result of not completing the cross-linking process on the surface of the membrane, which makes the cross-linking layer of the membrane, especially at higher pressures, the possibility of resistance to prevent the penetration of nitrogen molecules. lost and the permeability of nitrogen gas increases.

The findings showed that increasing the concentration of SiO_2 has the greatest effect on increasing membrane water flux. According to the other findings of this study, the effect of increasing the reaction time of the vapor phase carrier on reducing the selectivity reduction (drop trend) was at higher pressures. Although this measurement was not done due to the safety limitation of the equipment to measure the performance of membrane at pressures higher than 7 bar, but some graphs were presented regarding the state of penetrant which were a measure of the superiority of water component in the penetrant relative to gas component and the higher water component in the penetrant, is an indicator of the appropriateness of membrane's performance at its corresponding pressure, in such a way that the investigation of the difference trend of this two diagrams shows the possibility of using membrane at higher pressures.

According to the comparison between SiO_2 concentrations, compared to TIZ-V membrane, it is concluded that the increase in SiO_2 concentration has caused a slight decrease in gas permeability. An increase in SiO_2 provides the required amounts for the reaction with the existing amount of P-TiS, but its higher amounts will not cause more transverse connection of the network due to the lack of reaction with P-TiS, so if the results show that there is no major change in the amounts The permeability of the 4SiO2 membrane was due to the lack of major change in the transverse connection of the selective layer compared to the TiZ-V membrane.

Next, in order to increase the performance of membrane as much as possible, the sweeper gas was added from the inside of membrane. Increase in the sweeper gas, which increases the water concentration gradient and decreases the gas concentration gradient on the sides of membrane wall, increases the selectivity of membrane to the best level of 543. Finally, in order to investigate the stability of membrane in longer operation times, the test of increase the operation times was carried out and no change in relative humidity reduction data and insignificant changes in its selectivity indicate the stability of membrane performance with an increase in the dehumidification operation time.

5. ACKNOWLEDGMENTS

This research was done with the cooperation of the Southern Oil Regions Company. We are grateful to Dr. Seyyed Jalil Pour Mohamadian as an industrial consultant for this research.

6. REFERENCES

1. Zito PF, Brunetti A, Caravella A, Drioli E, Barbieri G. Water vapor permeation and its influence on gases through a zeolite-4A membrane. Journal of Membrane Science. 2019;574:154-63. https://doi.org/10.1016/j.memsci.2018.12.065
2. Delavar M, Nabian N. An investigation on the oxygen and nitrogen separation from air using carbonaceous adsorbents. J Eng Sci Technol. 2015;10(11):1394-403. https://doi.org/5e612106458515163551c787.pdf
3. Liang CZ, Chung T-S. Robust thin film composite PDMS/PAN hollow fiber membranes for water vapor removal from humid air and gases. Separation and Purification Technology. 2018;202:345-56. https://doi.org/10.1016/j.seppur.2018.03.005
4. Zhao B, Wang L-Y, Chung T-S. Enhanced membrane systems to harvest water and provide comfortable air via dehumidification & moisture condensation. Separation and Purification Technology. 2019;220:136-44. https://doi.org/10.1016/j.seppur.2019.03.034
5. Casado-Coterillo C, Fernández-Barquín A, Irabien A. Effect of humidity on CO2/N2 and CO2/CH4 separation using novel robust mixed matrix composite hollow fiber membranes: Experimental and model evaluation. Membranes. 2019;10(1):6. https://doi.org/10.3390/membranes10010006
6. Shirazian S, Ashrafizadeh SN. Synthesis of substrate-modified LTA zeolite membranes for dehydration of natural gas. Fuel. 2015;148:112-9. https://doi.org/10.1016/j.fuel.2015.01.086
7. Khulbe K, Matsuura T, Feng C, Ismail A. Recent development on the effect of water/moisture on the performance of zeolite membrane and MMMs containing zeolite for gas separation; review. RSC advances. 2016;6(49):42943-61. https://doi.org/10.1039/C6RA03007F
8. Scholes CA, Stevens GW, Kentish SE. Membrane gas separation applications in natural gas processing. Fuel. 2012;96:15-28. https://doi.org/10.1016/j.fuel.2011.12.074
9. Xu J, Zhang C, Ge T, Dai Y, Wang R. Performance study of sodium alginate-nonwoven fabric composite membranes for dehumidification. Applied Thermal Engineering. 2018;128:214-24. https://doi.org/10.1016/j.applthermaleng.2017.09.020
10. Salafi M, Asasian-Kolur N, Sharifian S, Ghadimi A. A flat-plate spiral-channeled membrane heat exchanger for methane dehumidification: Comparison of kraft paper and thin-film composite membrane. International Journal of Thermal Sciences. 2021;167:107046. https://doi.org/10.1016/j.ijthermalsci.2021.107046
11. Baker RW, Low BT. Gas separation membrane materials: a perspective. Macromolecules. 2014;47(20):6999-7013. https://doi.org/10.1021/ma501488s
12. Bui D, Vivekh P, Islam M, Chua K. Studying the characteristics and energy performance of a composite hollow membrane for air dehumidification. Applied Energy. 2022;306:118161. https://doi.org/10.1016/j.apenergy.2021.118161
13. Sun H, Chen B, Zhao G, Zhao Y, Yang M, Song Y. The enhancement effect of water-gas two-phase flow on depressurization process: Important for gas hydrate production. Applied Energy. 2020;276:115559. https://doi.org/10.1016/j.apenergy.2020.115559
14. Zhao J, Liu D, Yang M, Song Y. Analysis of heat transfer effects on gas production from methane hydrate by depressurization. International Journal of Heat and Mass Transfer. 2014;77:529-41. https://doi.org/10.1016/j.ijheatmasstransfer.2014.05.034
15. Zhao E, Hou J, Du Q, Liu Y, Ji Y, Bai Y. Numerical modeling of gas production from methane hydrate deposits using low-

frequency electrical heating assisted depressurization method. Fuel. 2021;290:120075. https://doi.org/10.1016/j.fuel.2020.120075

16. Zhao X, Qiu Z, Zhang Z, Zhang Y. Relationship between the gas hydrate suppression temperature and water activity in the presence of thermodynamic hydrate inhibitor. Fuel. 2020;264:116776. https://doi.org/10.1016/j.fuel.2019.116776
17. Norouzi N, Shiva N, Khajehpour H. Optimization of energy consumption in the process of dehumidification of natural gas. Biointerface Research in Applied Chemistry. 2021;11:14634-9. https://doi.org/10.33263/BRIAC116.1463414639
18. Fakharnezhad A, Masoumi S, Keshavarz P. Analysis of design parameter effects on gas dehumidification in hollow fiber membrane contactor: Theoretical and experimental study. Separation and Purification Technology. 2019;226:22-30. https://doi.org/10.1016/j.seppur.2019.05.073
19. Shadanfar H, Elhambakhsh A, Keshavarz P. Air dehumidification using various TEG based nano solvents in hollow fiber membrane contactors. Heat and Mass Transfer. 2021;57:1623-31. https://doi.org/10.1007/s00231-021-03057-2
20. Petukhov D, Komkova M, Brotsman V, Poyarkov A, Eliseev AA, Eliseev AA. Membrane condenser heat exchanger for conditioning of humid gases. Separation and Purification Technology. 2020;241:116697. https://doi.org/10.1016/j.seppur.2020.116697
21. Nikooei E, AuYeung N, Zhang X, Goulas K, Abbasi B, Dyall A, et al. Controlled dehumidification to extract clean water from a multicomponent gaseous mixture of organic contaminants. Journal of Water Process Engineering. 2021;43:102229. https://doi.org/10.1016/j.jwpe.2021.102229
22. Lotfi Z, Keshavarz P. High-efficiency water vapor absorption by tri-ethylene glycol combined with methyldiethanolamine. Separation and Purification Technology. 2021;270:118841. https://doi.org/10.1016/j.seppur.2021.118841
23. Zhao W, Lu H, Li C. Composite hollow fiber membrane dehumidification: A review on membrane module, moisture permeability and self-cleaning performance. International Journal of Heat and Mass Transfer. 2021;181:121832. https://doi.org/10.1016/j.ijheatmasstransfer.2021.121832
24. Madhumala M, Nagamani T, Sridhar S. Hollow Fiber Membrane Contactors for Dehumidification of Air. Hollow Fiber Membrane Contactors: CRC Press; 2020. p. 195-206.
25. Kaibollahi MM, Keshavarz P. Study of Water Vapor Absorption from Gas Using Tetra-Ethylene Glycol (TREG). 2022. https://doi.org/10.22034/ijche.2022.319813.1163
26. Kian Jon C, Islam MR, Kim Choon N, Shahzad MW, Kian Jon C, Islam MR, et al. Membrane Air Dehumidification. Advances in Air Conditioning Technologies: Improving Energy Efficiency. 2021:225-55. https://doi.org/10.1007/978-981-15-8477-0_6
27. Bucșă S, Năstase G, Șerban A, Ciocan M, Drughean L. Cooling and Dehumidification Systems Used in air Separation. International Multidisciplinary Scientific GeoConference: SGEM. 2019;19(6.1):139-44. https://doi.org/10.5593/sgem2019/6.1/S24.018
28. Amid M, Nabian N, Delavar M. Functionalized Halloysite Nanotubes and Graphene Oxide Nanosheets Fillers Incorporated in UF Membranes for Oil/Water Separation. International Journal of Engineering, Transactions A: Basics. 2023;36(7):1201-10. https://doi.org/10.5829/ije.2023.36.07a.01
29. Mohammadipour E, Nabian N, Delavar M. Novel PVC-melamine mixed matrix membranes for the Sirius Red removal from aqueous solutions: Experimental study and RSM modeling. Journal of Water Process Engineering. 2022;47:102752. https://doi.org/10.1016/j.jwpe.2022.102752
30. Shams K, Ahi H. Synthesis of 5A zeolite nanocrystals using kaolin via nanoemulsion-ultrasonic technique and study of its sorption using a known kerosene cut. Microporous and mesoporous materials. 2013;180:61-70. https://doi.org/10.1016/j.micromeso.2013.06.019
31. Matsukata M, Ogura M, Osaki T, Hari Prasad Rao PR, Nomura M, Kikuchi E. Conversion of dry gel to microporous crystals in gas phase. Topics in Catalysis. 1999;9:77-92. https://doi.org/10.1023/A:1019106421183
32. Liu X, Li Y, Chen B, Wang Y. Cleaner process for synthesis of zeolite MCM-22 by vapor-phase transport method. Asia-Pacific Journal of Chemical Engineering. 2009;4(5):607-11. https://doi.org/10.1002/apj.300

Persian Abstract

چکیده

هدف اصلی در این پژوهش استفاده از غشاهای نانو کامپوزیتی زئولیت TiO_2به منظورحذف رطوبت از گاز است. به همین جهت انتخاب و ساخت یک غشا TiZ-V معین به عنوان غشاء معیار، که حاصل براورد اولیه از یک غشا مناسب جهت نم زدایی از گاز بوده صورت گرفت و این غشا به عنوان معیاری برای سنجش اثر پارامترهای ساخت قرارگرفت . یافته ها نشان داد که افزایش غلظتSiO_2 دارای بیشترین اثر در افزایش شار آب غشا با توجه به تاثیر افزایش زمان واکنش حامل فاز بخار و کاهش روند افت گزینش پذیری در فشارهای بالاتر بوده است. همچنین آزمایشات تغیر رطوبت نسبی خوراک نشان دهنده کارایی بهتر غشا در رطوبت های نسبی پایین تر از 80٪ بوده است به نحوی که تحت شرایط رطوبت نسبی کمتر ، گزینش پذیری غشا افزایش یافته است. مورد مثبت دیگر یافت شده تغییر اندک کارایی گزینش پذیری غشا نسبت به رطوبت¬های نسبی مختلف است. این مورد نشان دهنده پایداری عملکرد غشا تحت شرایط مختلف رطوبت گاز خوراک است. سپس به منظور افزایش هرچه بیشتر عملکرد غشا، گاز سوییپر از داخل ممبران اضافه شد. افزایش گاز سوئیپر که باعث افزایش گرادیان غلظت آب و کاهش گرادیان غلظت گاز در طرفین دیواره غشا می شود، گزینش پذیری غشا را به بهترین سطح 543 افزایش می دهد.

IJE TRANSACTIONS B: Applications Vol. 37 No. 02, (February 2024) 377-386

International Journal of Engineering

Journal Homepage: www.ije.ir

Performances of Amphiphilic Glucomannan Produced by Combination Methods of Ultrasonication, Deacetylation, and Carboxymethylation Heterogeneously

D. H. Wardhani*[a], H. N. Ulya[a], A. Redondo[a], A. Riztian Nugraha[a], A. C. Kumoro[a], S. Susanti[b]

[a] Department of Chemical Engineering, Faculty of Engineering, Diponegoro University, Jl. Prof. Jacub Rais, Semarang, Central Java, Indonesia
[b] Department of Food Technology, Faculty of Animal and Agricultural Sciences, Diponegoro University, Jl. Prof. Soedarto, SH, Semarang, Central Java, Indonesia

PAPER INFO

Paper history:
Received 24 August 2023
Received in revised form 25 September 2023
Accepted 26 September 2023

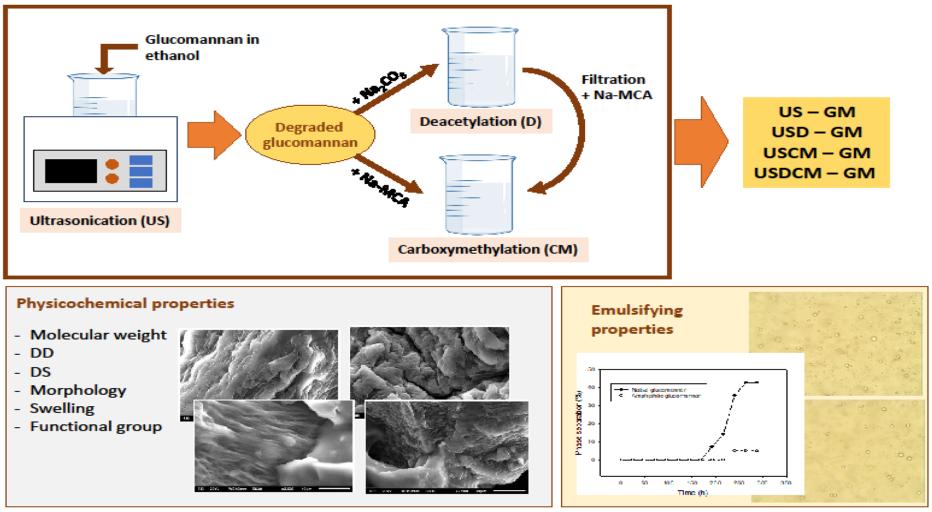
Keywords:
Emulsifier
Emulsion
Droplet
Molecular Weight
Viscosity

ABSTRACT

Glucomannan carboxymethylation is conducted to increase its hydrophobic properties and expand its interaction with hydrophobic compounds. However, glucomannan has high molecular weight and long polysaccharide chain which prevented the modification of its amphiphilic properties. This study aimed to examine the effect of molecular weight, deacetylation, and carboxymethylation on glucomannan properties. Performance of the modified glucomannan to stabilize oil in water (o/w) emulsion was also studied. Ultrasonication was applied to glucomannan at 40 kHz for 15-45 min to obtain various molecular weights. Sodium carbonate (Na_2CO_3) and sodium monochloroacetate were used as deacetylation and carboxymethylation agents, respectively. The results show that decrease molecular weight supported the deacetylation and the carboxymethylation process on attaching the hydrophobic groups to the glucomannan chain, hence, lowering the hydrophilic properties and swelling degree of glucomannan. Structural and morphological changes of glucomannan after modifications were confirmed from the IR spectra and SEM images. Excellent performance of the amphiphilic glucomannan on stabilizing o/w emulsion was observed as only ~5% phase separation occurred after 300 h of storage in ambient conditions. Hence, ultrasonication is proposed as a suitable preliminary treatment for amphiphilic glucomannan production.

***doi**: 10.5829/ije.2024.37.02b.13*

Graphical Abstract



---

[1] *Corresponding Author Email: dhwardhani@che.undip.ac.id (D. H. Wardhani)

Please cite this article as: Wardhani DH, Ulya HN, Redondo A, Riztian Nugraha A, Kumoro AC, Susanti S. Performances of Amphiphilic Glucomannan Produced by Combination Methods of Ultrasonication, Deacetylation, and Carboxymethylation Heterogeneously, International Journal of Engineering, Transactions B: Applications. 2024;37(02):377-86.

NOMENCLATURE

C	Concentration (g/l)	N	Normality concentration (N)
CM	Carboxymethylation	n	Group content
D	Deacetylation	t	Time (h or min or s)
DD	Degree of deacetylation	US	Ultrasonication
DS	Degree of substitution	V	Titrant volume (ml)
K	Mark-Houwink constant (ml/g)	**Greek Symbols**	
M	Molarity concentration (M)	η	Viscosity (Pa.s)
m	Material mass (g)	ω	Acetyl content
Mw	Molecular weight	α	Mark-Houwink constant

1. INTRODUCTION

Glucomannan, a polysaccharide extracted from *Amorphophallus sp*., is a neutral polysaccharide that is highly soluble in water. Apart from its high viscosity, glucomannan is also known for its gel-forming ability and low calories. Other than applied as a food ingredient, glucomannan is also used in the pharmaceutical, cosmetic, and chemical industries (1). However, some potential applications of glucomannan are limited by its high solubility and water absorption index which is up as high as 100 g water/g sample (2). The strong hydrophilicity of glucomannan led to poor water resistance and water barrier properties in its films which seriously hindered its practical applications (3).

For broader glucomannan applications, several methods have been developed to convert glucomannan into an amphiphilic compound which is capable of binding hydrophilic and lipophilic compounds (4). Molecular structure of material, such as the presence of long-chain hydrocarbon or short polymer chain, contributes on the amphiphilic properties (5). Hence, the amphiphilic glucomannan could be formed through a functionalization process by designing an attachment of functional groups or moieties to enhance its use (6). This amphiphilic polysaccharide has the potential to be applied in the plastics industry, especially in the production of biodegradable films, resins, and coatings for the pharmaceutical and food industries (5). The modified glucomannan could also be performed for example as a stabilizer in o/w emulsion and encapsulant oil-base bioactive compound (7).

It has been reported that carboxymethylation decreased the water-binding activity of starch and curdlan (8, 9). Carboxymethyl modification also successfully reduced solubility and water absorption of glucomannan, which further altered its surface properties (10, 11). In this modification, the hydrophilic groups of glucomannan, i.e. hydroxyl groups, are substituted with hydrophobic groups of monochloroacetic compounds (10). The modification consists of two steps of reactions: (1) deacetylation/alkalization and (2) carboxymethylation itself. Deacetylation or alkalization is performed to eliminate acetyl groups of glucomannan under an alkaline environment (12). This replacement reduces glucomannan steric hindrance and increases its intramolecular interaction; thus, leading to more active sites to be involved in the subsequent carboxymethylation reaction (13). The deacetylation was performed under an alkaline environment using NaOH, KOH, $Ca(OH)_2$, or Na_2CO_3 (14). Among those alkalis, Na_2CO_3 is safer to be consumed and does not disturb human body metabolism (15).

Xiao et al. (10) reported that various degrees of substitution (DS) using monochloroacetic acid that increase hydrophobic properties of glucomannan. In that work, the DS was obtained by varying reaction temperature, concentration of sodium, and reaction time. Moreover, the effectivity of the carboxymethylation process is influenced by the carboxymethyl reagents, temperature, duration, solvent, and polymer type (16).

Other than the presence of hydrophilic and hydrophobic groups, the length of each the groups also determines the performance of amphiphilic materials as surfactants. Day et al. (17) found an emulsion is more stable when using a surfactant with less content of the hydrophilic blocks.

Glucomannan has a molecular weight of up to 1.044×10^6 Da which makes it as one of the highest-viscosity polysaccharides (1%=15 Pa.s) (18). This condition and its ability to form gel prevent the glucomannan to interact with other molecules during modification. To obtain a high yield of modified glucomannan, decreasing the viscosity by lowering glucomannan concentration is not an option. Hence, glucomannan requires to be degraded prior to the amphiphilic modification. Degradation of glucomannan has been conducted through chemical, physical, and biological treatments to shorten the polymer chain, thus lowering its molecular weight (19). Among those methods, physical treatment was the preferable method to decrease the molecular weight due to safety reasons either for the process, the products or for the environment (20). Ultrasound, as one of the physical treatments, produced shear forces formed by the microbubble explosion, which is powerful enough to break the polysaccharide chain and reduce molecular weight of the compound (21, 22). Lower molecular weight of glucomannan has a lower viscosity which allows it to have more mobility hence higher chances of colliding with other molecules to pursue the reaction (23).

The significance of molecular weight on the emulsifying properties of carboxymethylated glucomannan remains limited. Hence, this research focused on studying the effect of molecular weight on the amphiphilic properties of modified glucomannan. Various molecular weights were obtained by controlling duration of the ultrasonication treatment. To modify the amphiphilic glucomannan, Na_2CO_3 was applied in deacetylation process while carboxymethylation was carried out using sodium monochloroacetate. The modifications were conducted subsequently under ethanol solution. The alter of glucomannan properties were observed, i.e. molecular weight, degree of deacetylation, degree of substitution, swelling degree, functional group, and surface morphology. The ability of native and modified glucomannan to stabilize o/w emulsion was also compared.

2. MATERIALS AND METHOD

2. 1. Materials Glucomannan of NOW Foods (Bloomingdale, Illinois, US) was used as it is, with 99.97% purity and weight average molecular weight (Mw) of 901,175 Da determined by the GPC method (24). Sodium monochloroacetate, Na_2CO_3, ethanol, phenolphthalein indicator, and other supporting materials were in pro-analyst standard.

2. 2. Glucomannan Modification In this study, glucomannan was treated physically and chemically. Firstly, glucomannan was physically degraded using ultrasonication to obtain various viscosity-average molecular weights (Mw). Subsequently, chemical modification to obtain the amphiphilic glucomannan was conducted using a combination of deacetylation and carboxymethylation. The detailed procedure is explained in the following sections.

2. 2. 1. Glucomanann Degradation Glucomannan degradation was conducted using sonication power in a heterogeneous system. Glucomannan (20 g) was placed in Erlenmeyer, together with ethanol solution (80%, 100 ml). The Erlenmeyer was closed using perforated lid and was put in an ultrasonic cleaner (Krisbow Indonesia, West Jakarta, Jakarta, Indonesia) which was filled with water and set at 40 kHz. The ultrasonication was carried out for 15, 30, and 45 min. The ultrasonicated glucomannan was separated from the ethanol by filtration and dried at room temperature for 24 h.

2. 2. 2. Preparation of an Amphiphilic Glucomannan The amphiphilic sample was prepared using deacetylation and carboxymethylation methods (14, 25). Deacetylation was carried out by dispersing the sonicated glucomannan (5 g) in ethanol solution (80%, 100 ml) and mixing it with 100 ml 0.4M Na_2CO_3 under constant stirring. After 24 h of the reaction, glucomannan was filtered and washed two times using 50 ml of 65% ethanol and 96% ethanol, subsequently. The deacetylated glucomannan was dried at room temperature for 24 h.

Carboxymethylation was conducted by mixing the sample (5 g) with sodium monochloroacetate (3.5 g) in ethanol solution (80%, 50 ml) at 60°C for 120 min. The glucomannan was filtered and washed using 100 ml of 96% ethanol. The amphiphilic glucomannan was then dried at room temperature for 24 h.

2. 3. Physicochemical and Stabilizing Properties The native and modified glucomannan were analyzed for their physicochemical properties, i.e. molecular weight, degree of deacetylation (DD), degree of substitution (DS), swelling power, functional group, and morphology. The emulsifying performance of native and modified glucomannan was also observed. The detailed procedure of the determinations is written below.

2. 3. 1. Viscosity-average Molecular Weight The viscosity-average molecular weight (Mw) of glucomannan was determined using the viscometry method (20). Flow time (t1) of the sample solution (0.1 – 0.5 g/l) in Cannon Fenske viscometer, size 100 (Schott AG, 102 Mainz, Rhineland-Palatinate, Germany) was recorded, compared to the flow time of distilled water as a solvent (t0). Relative viscosity (Ƞr), specific viscosity (Ƞsp), reduction viscosity (Ƞrd), and inherent viscosity (Ƞin) of the samples were calculated using Equation 1- 4, respectively.

$$\text{Ƞ}r = \frac{t1}{t0} \tag{1}$$

$$\text{Ƞsp} = \frac{t1-t0}{t0} = \text{Ƞr-1} \tag{2}$$

$$\text{Ƞrd} = \frac{\text{Ƞsp}}{C} \tag{3}$$

$$\text{Ƞin} = \frac{\ln(\text{Ƞrd})}{C} \tag{4}$$

The reduction viscosity and the inherent viscosity were then plotted in a graph. The intrinsic viscosity (Ƞ) was obtained by averaging the intercepts of both graphs. The average Mw of glucomannan was calculated using intrinsic viscosity value using Equation 5, with K=0.000506 ml/g and α=0.754 as Mark-Houwink constants (26).

$$\text{Mw} = \left(\frac{\text{Ƞ}}{K}\right)^{1/\alpha} \tag{5}$$

2. 3. 2. Degree of Deacetylation The back titration method was used to determine the degree of deacetylation (DD). Sample (1 g) was dispersed in ethanol solution (75%, 10 ml) at 50°C. After 30 min, 1 ml of 0.5 M KOH solution was added. Constant stirring was applied during the analysis. After 24 h, the mixture was titrated with 0.1 M HCl solution using phenolphthalein as an indicator. Titrant volume was recorded and used for calculating the DD using Equations 6-8.

$$\omega_0 = \frac{(V_2 - V_0) \times N_{HCl} \times M_{acetyl}}{m_s} \times 100\% \qquad (6)$$

$$\omega = \frac{(V_2 - V_1) \times N_{HCl} \times M_{acetyl}}{m_s} \times 100\% \qquad (7)$$

$$DD = \frac{\omega_0 - \omega}{\omega_0} \times 100\% \qquad (8)$$

where V_2, V_1, and V_o are the volume of HCl for the blank titration, the deacetylated glucomannan, and the native glucomannan, respectively. N_{HCl} is the concentration of HCl, Mw_{acetyl}=43 g/mol, and m_s is the mass of the sample to be titrated. Meanwhile, ω_0 and ω are the amounts of acetyl in the native glucomannan and in the deacetylated glucomannan, respectively.

2. 3. 3. Degree of Substitution The determination of the degree of substitution (DS) followed the method of Distantina et al. (27). Sample (1 g) was dispersed in a HCl solution (1.8 M, 10 ml) and stirred for 30 min. After filtration, the solid was washed using 96% ethanol solution and dried in ambient conditions. The dried glucomannan was dispersed in NaOH solution (0.2 M, 20 ml) and diluted to 100 ml with distilled water before being titrated using 0.05 M HCl with phenolphthalein as an indicator. The amount of substituted COOH (nCOOH) was calculated using Equation 9, with Vb and V are titrant volume of blank sample and glucomannan sample, respectively. Meanwhile, the DS was calculated using Equation 10.

$$n_{COOH} = (V_b - V).C_{HCl}.4 \qquad (9)$$

$$DS = \frac{162 \times n_{COOH}}{m_{ds} - 58 \times n_{COOH}} \qquad (10)$$

2. 3. 4. Swelling Swelling determination followed the method of Wardhani et al. (12). The glucomannan sample (0.1 g) was dispersed in distilled water (10 ml) at 60°C for 30 min prior to centrifugation at 4000 rpm for 20 min. The pellet was separated by filtration and dried. The weights of wet and dried pellets were recorded for swelling determination using Equation 11.

$$Swelling\ (\%) = \frac{weight\ of\ wet\ pellet}{weight\ of\ dried\ pellet} \times 100\% \qquad (11)$$

2. 3. 5. Functional Groups and Morphology The functional groups of the sample were observed from their infrared spectra which were obtained using Perkin Elmer Spotlight 200 FTIR (PerkinElmer Inc., Waltham, MA, USA) at the range of 4,000–400 cm^{-1} wavenumber. Meanwhile, the morphology of the samples was captured using an FEI Inspect S50 Scanning Electron Microscope (ThermoFisher Scientific, Waltham, Massachusetts, United States) at 5 kV after the gold-coating process.

2. 3. 6. Emulsion Stability and Optical Microscopy Ten milliliters of glucomannan emulsion was prepared by homogenizing 0.1 ml of fish oil and the glucomannan solution (1%) using a homogenizer (IKA RW homogenizer, Staufenim Breisgau, Germany) at 12000 rpm for 10 min in an ambient condition. The emulsion stability was determined by observing the height of the cream-serum separation (28). Meanwhile, the optical micrograph of the fresh emulsion was captured using a digital camera connected to a binocular microscope (Sinher XSZ-107BN, Ningbo Beilun Fangyuan Photoelectric Ltd., Zhejiang, China) at 40× objective lens magnification.

2. 4. Statistical Analysis The obtained data (except the Mw and instrumental data) were performed in triplicate. A one-way analysis of variance (ANOVA) was used for group comparison using Ms. Excel 2019. Data significances were set at $p<0.05$. The graphs were generated using Sigmaplot (Systat Software Inc., Chicago, IL, USA).

TABLE 1. Viscosity-average molecular weight (Mw) of native and modified glucomannan

Glucomannan treatments	US time (min)	Mw (g/mol)	Decrease of Mw (%)
Native	-	901,175	-
US[a]	15	145,711	83.83
	30	144,183	84.00
	45	142,821	84.15
(US+D)[b]	15	106,586	88.17
	30	93,853	89.58
	45	89,590	90.05
(US+CM)[c]	15	123,499	86.29
	30	119,779	86.70
	45	113,710	87.38
(US+D+CM)[d]	15	36,247	95.98
	30	20,730	97.70
	45	10,620	98.82

Note: a=ultrasonication,
b=ultrasonication+deacetylation, subsequently
c=ultrasonication+carboxymethylation, subsequently
d=ultrasonication+deacetylation+carboxymethylation, subsequently

3. RESULTS AND DISCUSSIONS

In this work, the amphiphilic glucomannan was prepared in two steps i.e., deacetylation and carboxymethylation. Prior to these treatments, glucomannan was sonicated to obtain various molecular weights. The effects of the molecular weight on the amphiphilic glucomannan properties after the modifications were studied.

3. 1. Glucomannan Degradation Effectiveness of ultrasonication on degrading glucomannan was represented by the decrease of viscosity-average Mw as shown in Table 1. Sonication for 45 min reduced viscosity-average Mw of glucomannan up to 84%. During sonication, the ultrasonic wave formed microbubbles surrounding the glucomannan which then exploded and produced cavitation effect on the glucomannan chain (20). Moreover, the waves also assisted the oscillation of polymer particles and caused more particle collisions (29). Both of those phenomena facilitated molecular degradation and resulted shorter chain of glucomannan, thus decreasing its viscosity-average Mw significantly, as also found by Wardhani et al. (30).

This ultrasonicated glucomannan was subsequently treated with deacetylation, carboxymethylation, or both combined processes. Deacetylation after sonication further reduced the viscosity-average Mw. Deacetylation lowered the viscosity-average Mw as the alkali group that replaced the acetyls of glucomannan had lower Mw (31). Hence, combination ultrasonication and deacetylation decreased the viscosity-average Mw of glucomannan further. Lower Mw of glucomannan after carboxymethylation process was also found by Wang et al. (25). However, the carboxymethyl group has higher Mw than that of the hydroxyl. Hence, the Mw of the carboxymethylation glucomannan was higher compared to that of the deacetylation.

Meanwhile, the viscosity-average Mw of the glucomannan reduced significantly by over 98% when deacetylation and carboxymethylation were applied serially. Replacement of the acetyl group during deacetylation removed steric hindrance of glucomannan to interact with other molecules. This suggested that the deacetylation facilitated the glucomannan to react with the sodium monochloroacetate, in the subsequent step. Moreover, the carboxymethylation produced alkali in the solution as the effect of carboxymethyl attachment, which caused further degradation of glucomannan chain (32).

Therefore, the highest decrease of viscosity-average Mw of glucomannan was found in combination treatment of ultrasonication, deacetylation, and carboxymethylation. Although each deacetylation or carboxymethylation gave further decrease of viscosity-average Mw after ultrasonication, however, the properties of both glucomannan modification could be different. Hence, the properties of the glucomannan should be considered while choosing the suitable degradation method.

3. 2. Degree of Deacetylation Deacetylation is conducted to substitute the acetyl group with the hydroxyl group which are more active in the carboxymethylation process (33). High Mw hindered the substitution reaction during deacetylation and led to have lower DD (Figure 1). Shorter the glucomannan chain improved glucomannan solubility which facilitated the deacetylation reaction as ionization was easier to occur (31). Moreover, lower molecular weight of glucomannan formed lower viscosity of glucomannan solution which eased the molecule mobility and collision (34). Hence, a lower molecular weight of glucomannan is preferable to increase DD.

3. 3. Degree of Substitution Ultrasonication and deacetylation, as preliminary treatments prior to the glucomannan carboxymethylation, improved the level of DS (Figure 1). Ultrasonication had a role in lowering glucomannan Mw (Table 1). Xiao et al. (10) and Arifan and Primartu (35) applied deacetylation prior to carboxymethylation to obtain the effective substitution. As mentioned in the earlier section, lowering molecular weight increased the glucomannan solubility and decreased its viscosity, thus leading to ease the glucomannan reaction. High Mw tended to have rigid structures which hindered the carboxymethylation agent from interacting and substituting the hydroxyl groups (36). Deacetylation not only removed the acetyl content of glucomannan but also provided more hydroxyl groups for carboxymethylation substitutions (37). Carboxymethyl groups of sodium monochloroacetate reacted with alkoxides of glucomannan produced after deacetylation/alkalization process (16). Therefore, the

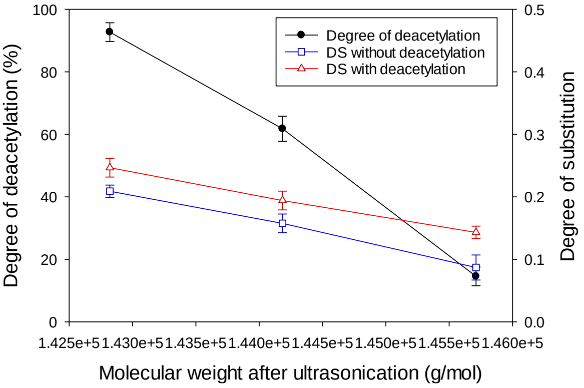


Figure 1. The effect of viscosity-average molecular weight of glucomannan on degree of deacetylation (DD) and degree of substitution (DS)

deacetylation promoted carboxymethylation and increased the DS. Lower Mw due to sonication had indirect effect of increasing degree of substitution through intensifying deacetylation process. Hence, to maximise DS, it is suggested to decrease Mw of polysaccharides and conduct deacetylation process prior to carboxymethylation modification.

3. 4. Swelling Swelling determination represents the ability of glucomannan molecules to expand in aqueous solution. Other than the structural properties, swelling degree of material is also influenced by its charged ions and solvent type (38). Figure 2 shows that ultrasonication improve the swelling of glucomannan. Wang et al. (39) reported that ultrasonication destroyed its crystalline structure which reduced the water-binding ability. Wu et al. (40) found that ultrasonic cavitation also lowered the gel-forming ability of glucomannan.

Deacetylation and carboxymethylation caused further decrease of sweling power. Li et al. (41) explained that water penetration could cause glucomannan swelling through three steps: (i) loosening the lamellar structure of glucomannan, (ii) breaking the intermolecular forces between the lamellar, and (iii) decomposition of amorphous region and severing the crystalline region. Those steps increased the area for water to penetrate. Deacetylation removed the steric hindrance among the glucomannan molecules. This condition allowed glucomannan to interact and reduced the number of water sorption sites (42) resulted lower swelling ability. Carboxymethylation reinforced the glucomannan structure by generating negative charge on the carboxymethylated glucomannan. This negative charged allowed it to interact with proton-hydroxyl group of other glucomannan molecules (36). Decrease of water absorption after glucomannan carboxymethylation was also reported by Xiao et al. (10). Among all glucomannan treatments in this work, the combination of ultrasonication-deacetylation-carboxymethylation had the most significant effect on decreasing the swelling, hence, this series of reactions gave synergetic work to improve the carboxymethylation reaction and reduced the swelling power.

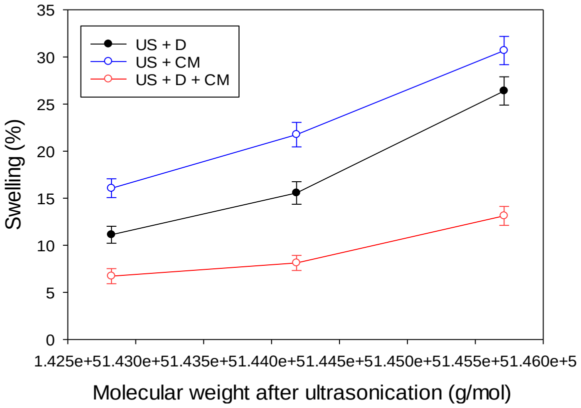
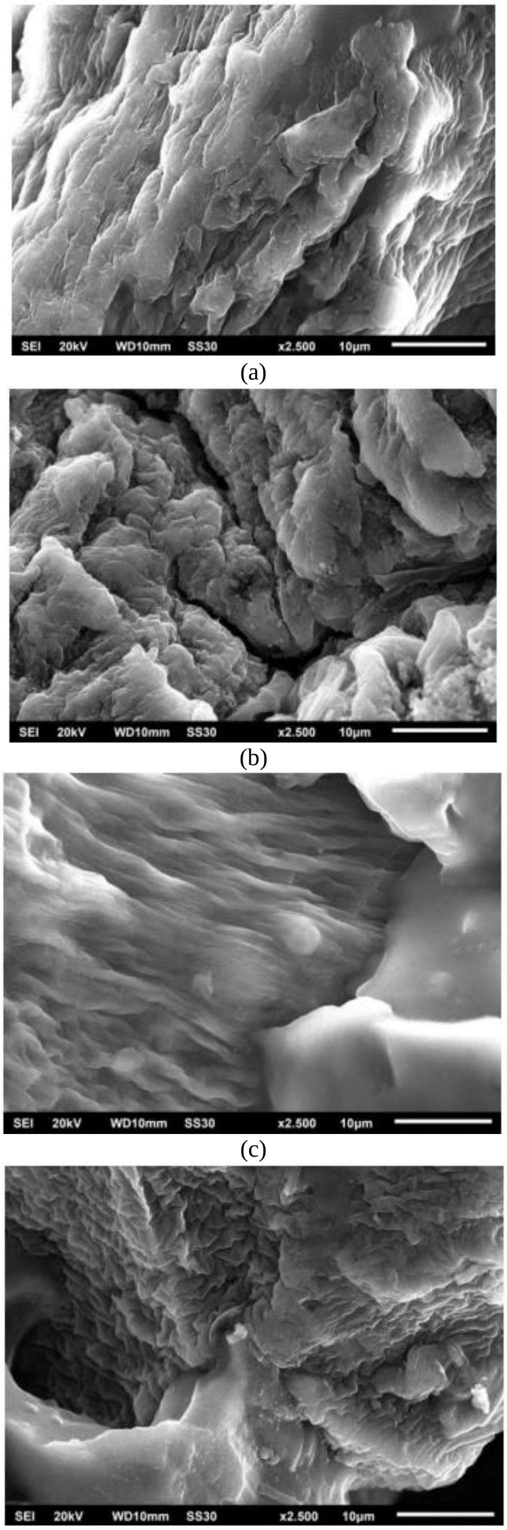


Figure 2. Effect of ultrasonication (US), deacetylation (D), and carboxymethylation (CM) on the swelling of modified glucomannan

3. 5. Morphology The surface morphology of the modified glucomannans is shown in Figure 3. Ultrasonicated glucomannan has a rough surface and irregular form. After deacetylation, more wrinkles are observed with deeper cleavages. Different surface morphology is shown by carboxymethylated glucomannan, which has smaller wrinkles. The surface morphology changes after deacetylation and carboxymethylation process were also found by Liu et al. (43) and Xiao et al. (10). The SEM images indicated that chemical modification, i.e. deacetylation and carboxymethylation, also attacked the physical structure of glucomannan.

3. 6. Functional Groups Generally, the modifications applied to the glucomannan did not change the forms of functional groups but their transmittance intensities (Figure 4). Native glucomannan shows its characteristic peaks at ~3325 cm^{-1} (O-H) (44), ~2825 cm^{-1} (C-H), ~1750 cm^{-1} (C=O of the acetyl), ~1250 cm^{-1} (C-O), and ~1000 cm^{-1} (C=O-C). Meanwhile, the mannose group of glucomannan was identified at ~800 cm^{-1}. Ultrasonication and deacetylation increased the peak of O-H group. Ultrasonication increased the hydroxyl content of glucomannan as the effect of its chain-breaking (30), while the increase after deacetylation indicated that the hydroxyl group attachment occurred during the reaction. Jin et al. (45) also found an increasing O-H group intensity after deacetylation, which also showed an increase in intermolecular hydrogen bonds (46).

Interestingly, almost all peaks of the sonicated-carboxymethylated sample showed the lowest intensity of transmittance except on ~1400 cm^{-1} which is related to -CH$_2$- group. Deacetylation removed C=O of acetyl from the sonicated sample at ~1750 cm^{-1} due to alkali treatment. However, carboxymethylation reintroduced this group back to the sample, resulting in lower transmitance. The peaks of band at ~1250 and ~1017 cm^{-1}, which indicated the group of C-O and C=O-C, were intensified after carboxymethylation process. Stronger after carboxymethylation was also found by Wang et al. (25) and Chen et al. (47). Moreover, number of intramolecular hydrogen binding were found to increase after the carboxymethylation, as described by reducing the transmittance at ~1640 cm^{-1} (25). Therefore, the IR spectra showed that lower Mw promoted the carboxymethylation process, even without glucomannan deacetylation.



(a)

(b)

(c)

(d)

Figure 3. Surface morphology at 2,500× magnification of glucomannan after (a) ultrasonication, (b) ultrasonication-deacetylation, (c) ultrasonication-carboxymethylation, (d) ultrasonication-deacetylation-carboxymethylation

3. 7. Emulsion Stability

Performance of native and amphiphilic glucomannan on stabilizing o/w emulsion was determined by observing development of phase separation of the emulsion during 300 h storage at room temperature. The highest DS sample (DS=0.25) was selected to represent the amphiphilic glucomannan. The emulsion of the native glucomannan was separated after 180 h of storage (Figure 5). Meanwhile, the amphiphilic emulsion started to separate after 240 h. By the end of storage duration (300 h), the phase separation of the native and the amphiphilic emulsions was 42.88% and 5.15%, respectively.

The emulsion stability was also supported by the micrograph of glucomannan-fish oil droplets (Figure 6). After 300 h, more smaller droplets were observed in the amphiphilic glucomannan emulsion compared to those of

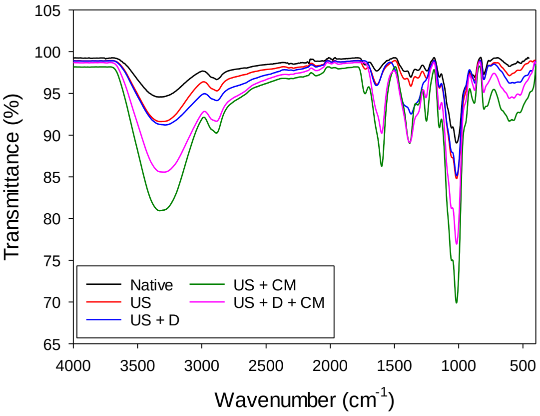


Figure 4. IR spectra of native glucomannan and modified glucomannan using ultrasonication (US), deacetylation (D), and carboxymethylation (CM)

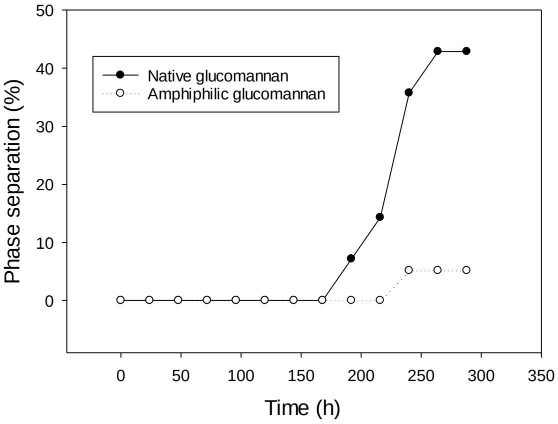
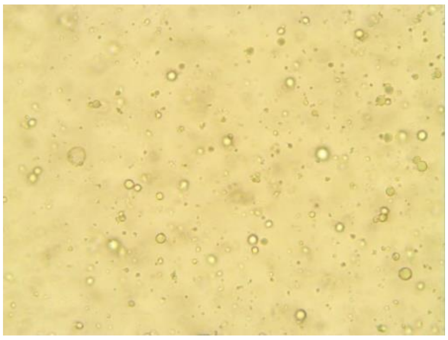
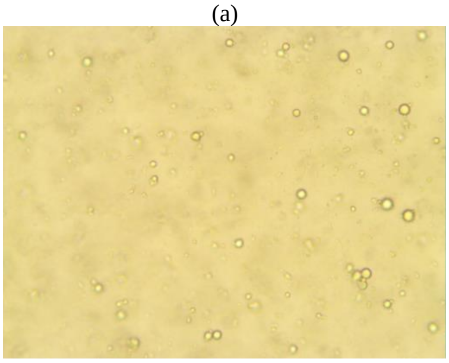


Figure 5. Phase separation of fish oil emulsion stabilized using native and amphiphilic glucomannan



(a)



(b)

Figure 6. Optical micrograph of fish oil emulsion right after homogenization stabilized using (a) native and (b) amphiphilic glucomannan

the native one. Droplet size of the emulsion was known as one of the emulsion stability indicators as smaller droplets slowed down the rate of droplet coalescence (48). Hence, the amphiphilic glucomannan could maintain the oil droplets in glucomannan solution and stablelize the emulsion. Other than droplet size, o/w emulsion stability was also influenced by viscosity, pH, ionic content, mixing variables, and interaction between dispersed and continuous phases (49).

As shown in Figure 4, the amphiphilic glucomannan has lower transmittance of carboxylate and methylene groups which have a hydrophobic property (50, 51). These hydrophobic groups are allowed to bond with fish oil during dissolving the oil in water. Cai et al. (52) also found a stable emulsion when used carboxymethyl starch in pH variation. This result suggested that serial treatment using ultrasonication, deacetylation, and carboxymethylation succeeded in improving the hydrophobic properties of glucomannan, thus, enhance its emulsifying properties. However, apart of the hydrophobic modification of glucomannan and droplet size, other factors should be examined in the future to precise the influencing factors on stabilized emulsion.

4. CONCLUSION

Natural emulsifier produced by native natural ingredient is lacked for its performance on stabilizing o/w emulsion. Decreasing the length of hydrophilic chain is known to increase the emulsifying properties of surfactants. In this study, amphiphilic glucomannan has been produced using carboxymethylation method which is supported by ultrasonication to initially degrade the glucomannan to obtain various Mw. Ultrasonication, deacetylation, and carboxymethylation performed synergetic work on decreasing the Mw of glucomannan and swelling power, while enhancing the DD of deacetylationand also the value of DS of carboxymethylation. The modifications changed the morphology and intensities of the transmitance peaks as shown in IR spectra of glucomannan. The amphiphilic glucomannan was successfully stabilized fish oil emulsion with ~5% phase separation after 300 h. This reduced separation exhibited the preferential stabilizing properties of amphiphilic glucomannan which is potential for food application. This study increased the probability to use natural surfactant for emulsion stabilization and encapsulation. The efficacy of the amphiphilic glucomannan to stabilize the emulsion of real foods and its impact on the food properties become a target for future study.

5. ACKNOWLEDGMENT

This study was funded by Directorate of Research and Community Service, Directorate General of Higher Education, Ministry of Research, Technology and Higher Education of the Republic of Indonesia through *Penelitan Fundamental*-2023 (No. 449A-40/UN7.D2/PP/VI/2023).

6. REFERENCES

1. Srzednicki G, Borompichaichartkul C. Konjac glucomannan: Production, processing, and functional applications: CRC Press; 2020.
2. Koroskenyi B, McCarthy SP. Synthesis of acetylated konjac glucomannan and effect of degree of acetylation on water absorbency. Biomacromolecules. 2001;2(3):824-6. https://doi.org/10.1021/bm010014c
3. Wu K, Zhu Q, Qian H, Xiao M, Corke H, Nishinari K, et al. Controllable hydrophilicity-hydrophobicity and related properties of konjac glucomannan and ethyl cellulose composite films. Food Hydrocolloids. 2018;79:301-9. https://doi.org/10.1016/j.foodhyd.2017.12.034
4. Dan N. Core–shell drug carriers: liposomes, polymersomes, and niosomes. Nanostructures for drug delivery: Elsevier; 2017. p. 63-105.
5. Cunha AG, Gandini A. Turning polysaccharides into hydrophobic materials: a critical review. Part 2. Hemicelluloses, chitin/chitosan, starch, pectin and alginates. Cellulose. 2010;17:1045-65. https://doi.org/10.1007/s10570-010-9435-5
6. Nworie FS, Mgbemena N, Ike-Amadi A, Ebunoha J. Functionalized Biochars for Enhanced Removal of Heavy Metals from Aqueous Solutions: Mechanism and Future Industrial Prospects. Journal of Human, Earth, and Future. 2022;3(3):377-95. https://doi.org/10.28991/HEF-2022-03-03-09

7. Zhang Y-q, Xie B-j, Gan X. Advance in the applications of konjac glucomannan and its derivatives. Carbohydrate polymers. 2005;60(1):27-31. https://doi.org/10.1016/j.carbpol.2004.11.003

8. Cova A, Sandoval AJ, Balsamo V, Müller AJ. The effect of hydrophobic modifications on the adsorption isotherms of cassava starch. Carbohydrate Polymers. 2010;81(3):660-7. https://doi.org/10.1016/j.carbpol.2010.03.028

9. Jin Y, Zhang H, Yin Y, Nishinari K. Comparison of curdlan and its carboxymethylated derivative by means of Rheology, DSC, and AFM. Carbohydrate Research. 2006;341(1):90-9. https://doi.org/10.1016/j.carres.2005.11.003

10. Xiao M, Dai S, Wang L, Ni X, Yan W, Fang Y, et al. Carboxymethyl modification of konjac glucomannan affects water binding properties. Carbohydrate Polymers. 2015;130:1-8. https://doi.org/10.1016/j.carbpol.2015.05.001

11. Nematov D, Kholmurodov K, Husenzoda M, Lyubchyk A, Burhonzoda A. Molecular Adsorption of H2O on TiO2 and TiO2: Y Surfaces. Journal of Human, Earth, and Future. 2022;3(2):213-22. https://doi.org/10.28991/HEF-2022-03-02-07

12. Wardhani DH, Nugroho F, Aryanti N, Prasetyaningrum A. Simultaneous effect of temperature and time of deacetylation on physicochemical properties of glucomannan. ASEAN Journal of Chemical Engineering. 2018;18(1):1-8.

13. Qiao D, Lu J, Shi W, Li H, Zhang L, Jiang F, et al. Deacetylation enhances the properties of konjac glucomannan/agar composites. Carbohydrate Polymers. 2022;276:118776. https://doi.org/10.1016/j.carbpol.2021.118776

14. Wardhani DH, Ulya HN, Nugroho F, Sasongko SB, Hadiyarto A, Nugroho A. Vitamin C encapsulation by a gelation method using deacetylated glucomannan as a matrix. Journal of King Saud University-Science. 2020;32(7):2924-30. https://doi.org/10.1016/j.jksus.2020.07.010

15. Du X, Li J, Chen J, Li B. Effect of degree of deacetylation on physicochemical and gelation properties of konjac glucomannan. Food Research International. 2012;46(1):270-8. https://doi.org/10.1016/j.foodres.2011.12.015

16. Aprilia V, Murdiati A, Hastuti P, Harmayani E. Carboxymethylation of glucomannan from porang tuber (Amorphophallus oncophyllus) and the physicochemical properties of the product. Pakistan Journal of Nutrition. 2017;16(11):835-42. https://doi.org/10.3923/pjn.2017.835.842

17. Day RA, Estabrook DA, Wu C, Chapman JO, Togle AJ, Sletten EM. Systematic study of perfluorocarbon nanoemulsions stabilized by polymer amphiphiles. ACS applied materials & interfaces. 2020;12(35):38887-98. https://doi.org/10.1021/acsami.0c07206

18. Li B, Xie B. Synthesis and characterization of konjac glucomannan/poly (vinyl alcohol) interpenetrating polymer networks. Journal of applied polymer science. 2004;93(6):2775-80. https://doi.org/10.1002/app.20769

19. Li J, Li B, Geng P, Song A-X, Wu J-Y. Ultrasonic degradation kinetics and rheological profiles of a food polysaccharide (konjac glucomannan) in water. Food Hydrocolloids. 2017;70:14-9. https://doi.org/10.1016/j.foodhyd.2017.03.022

20. Wardhani DH, Aryanti N, Aziz A, Firdhaus RA, Ulya HN. Ultrasonic degradation of alginate: A matrix for iron encapsulation using gelation. Food Bioscience. 2021;41:100803. https://doi.org/10.1016/j.fbio.2020.100803

21. Feng L, Cao Y, Xu D, Wang S, Zhang J. Molecular weight distribution, rheological property and structural changes of sodium alginate induced by ultrasound. Ultrasonics sonochemistry. 2017;34:609-15. https://doi.org/10.1016/j.ultsonch.2016.06.038

22. Low S, Tan M, Chin N, Tan K. Comparative Studies on Ultrasound Pre-treated Peanut Husk Powder and Ultrasound Simultaneous Process on Heavy Metal Adsorption. International Journal of Engineering. 2018;31(8):1334-40. https://doi.org/10.5829/ije.2018.31.08b.23

23. Wu Z, Song X. Carboxymethylation of γ-irradiated starch. Journal of applied polymer science. 2006;101(4):2210-5. https://doi.org/10.1002/app.23479

24. Wardhani DH, Cahyono H, Ulya HN, Kumoro AC, Anam K, Vázquez JA. Spray-dryer feed preparation: Enzymatic degradation of glucomannan for iron nanoencapsulation. 2022. https://doi.org/10.3934/agrfood.2022042

25. Wang M, He W, Wang S, Song X. Carboxymethylated glucomannan as paper strengthening agent. Carbohydrate polymers. 2015;125:334-9. https://doi.org/10.1016/j.carbpol.2015.02.060

26. Ratcliffe I, Williams PA, Viebke C, Meadows J. Physicochemical characterization of konjac glucomannan. Biomacromolecules. 2005;6(4):1977-86. https://doi.org/10.1021/bm0492226

27. Distantina S, Kaavessina M, Fadilah F, Putrie AS, Novianti I, editors. Carboxymethyl konjac glucomannan from konjac flour: The effect of media and temperature on carboxymethylation rate. AIP Conference Proceedings; 2018: AIP Publishing.

28. Huang X, Kakuda Y, Cui W. Hydrocolloids in emulsions: particle size distribution and interfacial activity. Food hydrocolloids. 2001;15(4-6):533-42. https://doi.org/10.1016/S0268-005X(01)00091-1

29. Li J, Cai J, Fan L. Effect of sonolysis on kinetics and physicochemical properties of treated chitosan. Journal of Applied Polymer Science. 2008;109(4):2417-25. https://doi.org/10.1002/app.28339

30. Wardhani DH, Ulya HN, Maulana I, Salsabila S, Kumoro AC, Vázquez JA. Analyzing the characteristics of degraded glucomannan of Amorphophallus oncophyllus using hydrogen peroxide and ultrasonication. 2023. https://doi.org/10.3934/agrfood.2023031

31. Ulya HN, Wardhani DH, Aryanti N, Pangestuti DR, editors. Deacetylation of glucomannan of Amorphophallus oncophillus using NaOH and its properties as iron excipient by gelation in ethanol. AIP Conference Proceedings; 2019: AIP Publishing.

32. El-Sheikh MA. Carboxymethylation of maize starch at mild conditions. Carbohydrate Polymers. 2010;79(4):875-81. https://doi.org/10.1016/j.carbpol.2009.10.013

33. Fadilah F, Distantina S, Kaavessina M, Wijayanti ST, Andayani R, editors. Study on the carboxymethylation of glucomannan from porang. AIP Conference Proceedings; 2018: AIP Publishing.

34. Guo L, Yokoyama W, Chen L, Liu F, Chen M, Zhong F. Characterization and physicochemical properties analysis of konjac glucomannan: Implications for structure-properties relationships. Food Hydrocolloids. 2021;120:106818. https://doi.org/10.1016/j.foodhyd.2021.106818

35. Arifan F, Primartu S. Physical and chemical characteristics of carboxymethyl cellulose from Kepok banana (Musa paradisiaca formatypica) and corncobs (Zea mays) modified alkalization-carboxymethylation process in ice cream products. Food Research. 2023;7(2):31-6. https://doi.org/10.26656/fr.2017.7(2).355

36. Wu J, Zhong F, Li Y, Shoemaker C, Xia W. Preparation and characterization of pullulan–chitosan and pullulan–carboxymethyl chitosan blended films. Food Hydrocolloids. 2013;30(1):82-91. https://doi.org/10.1016/j.foodhyd.2012.04.002

37. Li J, Ye T, Wu X, Chen J, Wang S, Lin L, et al. Preparation and characterization of heterogeneous deacetylated konjac glucomannan. Food Hydrocolloids. 2014;40:9-15. https://doi.org/10.1016/j.foodhyd.2014.02.001

38. Gani BA, Asmah N, Soraya C, Syafriza D, Rezeki S, Nazar M, et al. Characteristics and Antibacterial Properties of Film Membrane of Chitosan-Resveratrol for Wound Dressing. Emerging Science Journal. 2023;7(3):821-42. https://doi.org/10.28991/ESJ-2023-07-03-012
39. Wang M, Wu Y, Liu Y, Ouyang J. Effect of ultrasonic and microwave dual-treatment on the physicochemical properties of chestnut starch. Polymers. 2020;12(8):1718. https://doi.org/10.3390/POLYM12081718
40. Wu H, Tong L, Wang Y, Yan H, Sun Z. Bibliometric analysis of global research trends on ultrasound microbubble: a quickly developing field. Frontiers in Pharmacology. 2021;12:646626. https://doi.org/10.3389/fphar.2021.646626
41. Li L, Ruan H, Ma L-l, Wang W, Zhou P, He G-q. Study on swelling model and thermodynamic structure of native konjac glucomannan. Journal of Zhejiang University SCIENCE B. 2009;10:273-9. https://doi.org/10.1631/jzus.B0820221
42. Huang Y-C, Chu H-W, Huang C-C, Wu W-C, Tsai J-S. Alkali-treated konjac glucomannan film as a novel wound dressing. Carbohydrate Polymers. 2015;117:778-87. https://doi.org/10.1016/j.carbpol.2014.10.047
43. Liu X, Gan J, Nirasawa S, Tatsumi E, Yin L, Cheng Y. Effects of sodium carbonate and potassium carbonate on colloidal properties and molecular characteristics of konjac glucomannan hydrogels. International journal of biological macromolecules. 2018;117:863-9. https://doi.org/10.1016/j.ijbiomac.2018.05.176
44. Yusmaniar Y, Julio E, Rahman A, Yudanto SD, Susetyo FB. Synthesis of Polyvinyl Alcohol-Chitosan Composite Film using Nanocellulose from Coconut Fibers (Cocos nucifera). International Journal of Engineering. 2023;36(11):1993-2003. https://doi.org/10.5829/ije.2023.36.11b.05
45. Jin W, Song R, Xu W, Wang Y, Li J, Shah BR, et al. Analysis of deacetylated konjac glucomannan and xanthan gum phase separation by film forming. Food Hydrocolloids. 2015;48:320-6. https://doi.org/10.1016/j.foodhyd.2015.02.007
46. Solo-de-Zaldívar B, Tovar C, Borderías A, Herranz B. Effect of deacetylation on the glucomannan gelation process for making restructured seafood products. Food Hydrocolloids. 2014;35:59-68. https://doi.org/10.1016/j.foodhyd.2013.04.009
47. Chen L, Du Y, Tian Z, Sun L. Effect of the degree of deacetylation and the substitution of carboxymethyl chitosan on its aggregation behavior. Journal of Polymer Science Part B: Polymer Physics. 2005;43(3):296-305. https://doi.org/10.1021/bm010014c
48. Wardhani DH, Ulya HN, Kumoro AC, Aryanti N. Partial hydrophobic modification of alginate using dodecenyl succinic anhydride for fish oil encapsulation. Acta Scientiarum Polonorum Technologia Alimentaria. 2022;21(4):369-78. https://doi.org/10.17306/J.AFS.2022.1077
49. Maphosa Y, Jideani VA. Factors affecting the stability of emulsions stabilised by biopolymers. Science and technology behind Nanoemulsions. 2018;65. https://doi.org/10.5772/intechopen.75308
50. Spychaj T, Wilpiszewska K, Zdanowicz M. Medium and high substituted carboxymethyl starch: synthesis, characterization and application. Starch-Stärke. 2013;65(1-2):22-33. https://doi.org/10.1002/star.201200159
51. Yanli W, Wenyuan G, Xia L. Carboxymethyl Chinese yam starch: synthesis, characterization, and influence of reaction parameters. Carbohydrate Research. 2009;344(13):1764-9. https://doi.org/10.1016/j.carres.2009.06.014
52. Cai X, Wang Y, Du X, Xing X, Zhu G. Stability of pH-responsive Pickering emulsion stabilized by carboxymethyl starch/xanthan gum combinations. Food Hydrocolloids. 2020;109:106093. https://doi.org/10.1016/j.foodhyd.2020.106093

Persian Abstract

چکیده

کربوکسی متیلاسیون گلوکومانان برای افزایش خواص آبگریز آن و گسترش تعامل آن با ترکیبات آبگریز انجام می شود .با این حال، گلوکومانان دارای وزن مولکولی بالا و یک زنجیره پلی ساکارید طولانی است که از تغییر خواص آمفیفیلیک آن جلوگیری می کند .این مطالعه با هدف بررسی اثر وزن مولکولی، استیل زدایی و کربوکسی متیلاسیون بر خواص گلوکومانان انجام شد .عملکرد گلوکومانان اصلاح‌شده برای تثبیت امولسیون‌های روغن در آب (o/w) نیز مورد مطالعه قرار گرفت .برای به دست آوردن وزن های مولکولی مختلف، فراصوت به گلوکومانان با فرکانس 40 کیلوهرتز به مدت 15-45 دقیقه اعمال شد .کربنات سدیم (Na_2CO_3) و مونوکلرواستات سدیم به ترتیب به عنوان عوامل استیل زدایی و کربوکسی متیلاسیون استفاده شد .نتایج نشان می‌دهد که کاهش وزن مولکولی از فرآیند استیل‌زدایی و کربوکسی متیلاسیون در اتصال گروه‌های آبگریز به زنجیره گلوکومانان پشتیبانی می‌کند، بنابراین، خواص آبدوستی و درجه تورم گلوکومانان را کاهش می‌دهد .تغییرات ساختاری و مورفولوژیکی گلوکومانان پس از تغییرات از طریق طیف IR و تصاویر SEM تایید شد .عملکرد عالی گلوکومانان آمفیفیلیک در تثبیت امولسیون o/w مشاهده شد زیرا تنها 5 درصد جداسازی فاز پس از 300 ساعت ذخیره سازی در شرایط محیطی رخ داد .از این رو، سونوگرافی به عنوان یک درمان اولیه مناسب برای تولید گلوکومانان آمفیفیلیک پیشنهاد می شود.a

IJE TRANSACTIONS B: Applications Vol. 37 No. 02, (February 2024) 387-401

International Journal of Engineering

Journal Homepage: www.ije.ir

Enhancing Fault Detection in Image Analysis: A Combined Wavelet-Fourier Technique for Advancing Manufacturing Quality Control

Z. Khodadadi[a], M. S. Owlia*[a], A. Amiri[b]

[a] *Department of Industrial Engineering, Yazd University, Yazd, Iran*
[b] *Department of Industrial Engineering, Shahed University, Tehran, Iran*

P A P E R I N F O

Paper history:
Received 27 August 2023
Received in revised form 27 September 2023
Accepted 02 October 2023

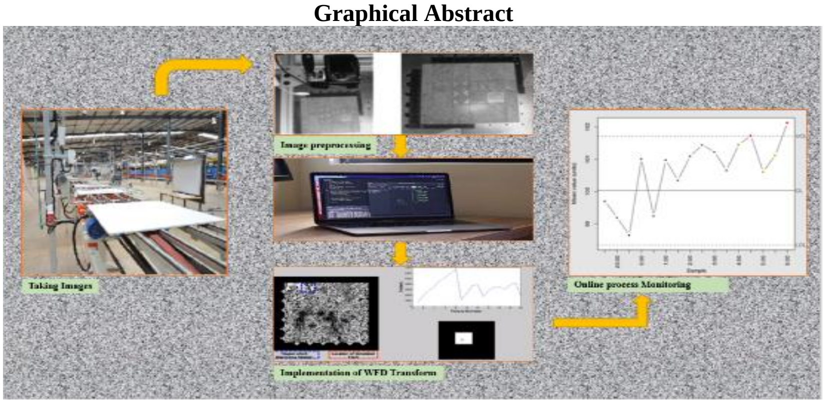
Keywords:
Quality Performance
Imaging Technologies
Feature Extraction
Statistical Process Control
Online Process Monitoring
Wavelet-Fourier Method
Fault Detection and Localization

A B S T R A C T

This study focuses on utilizing image data for statistical process control and improving quality monitoring in manufacturing and service systems. The effectiveness of individual and combined feature extraction methods is evaluated, with the Wavelet-Fourier approach identified as the most suitable. The proposed method not only identifies image processing issues but also provides valuable information for estimating change points, fault locations, and fault sizes. This enables the resolution and prediction of faults, leading to cost and time savings in production. To perform evaluation of the proposed method, an image from a tile production line is subjected to Wavelet transform, followed by Fourier transform on the obtained coefficients. The results demonstrate the superiority of the Wavelet-Fourier method over individual methods such as Fourier transform and Wavelet transform. The proposed method exhibits comparable or improved performance in fault detection and localization compared to similar research. This study highlights the potential of utilizing image data for statistical process control and quality monitoring, offering a comprehensive solution for fault detection and analysis. The findings contribute to advancements in image processing techniques and have practical implications for enhancing quality monitoring in various industries. By leveraging image data, manufacturers can make informed decisions, enhance process performance, and improve overall product quality.

doi*: 10.5829/ije.2024.37.02b.14*

Graphical Abstract



1. INTRODUCTION

In contemporary manufacturing and service industries, ensuring product reliability and customer satisfaction hinges on the successful implementation of robust quality control measures throughout production processes. The deployment of image analysis tools has proven invaluable in the detection of faults and defects across a wide spectrum of industries (1).

---

*Corresponding Author Email: *owliams@yazd.ac.ir* (M. S. Owlia)

Please cite this article as: Khodadadi Z, Owlia MS, Amiri A. Enhancing Fault Detection in Image Analysis: A Combined Wavelet-Fourier Technique for Advancing Manufacturing Quality Control, International Journal of Engineering, Transactions B: Applications. 2024;37(02):387-401.

The direct utilization of images within the domain of statistical process control, while promising, introduces its own set of challenges. These challenges encompass increased computational complexity and the vulnerability of features to variations. As the number of features derived from images grows, so does the computational burden, potentially impeding the model's ability to discern intricate relationships within high-dimensional data. Furthermore, the raw data extracted from images can be highly sensitive to even minor parameter fluctuations, affecting pixel values and potentially undermining the model's accuracy and reliability.

Image processing entails the conversion of an input image into a measurable or descriptive index as output, relying on the computation of a sequence of general or specific characteristics. In this process, the initial step involves capturing an image of the subject or object of interest. Subsequent analysis of these images necessitates the use of various computer facilities and software tools. Finally, by crafting suitable algorithms, a model is developed with the capacity to predict the current or future state of the target based on the extracted image features.

Descriptors, serving as representations of image quality aligned with human perception, facilitate effective image retrieval by similarity. While the Fourier transform has been widely employed for image description, it falls short in terms of resolution properties. Conversely, the Fourier descriptor demonstrates enhanced accuracy, efficiency, and performance in recovery and classification. This method allows for precise matching, ease of implementation, and sensitivity to low levels of noise. An increasingly popular alternative is the wavelet transform, which amalgamates the time and frequency domains, combining the advantages of both approaches. The wavelet transform dissects signals into fundamental functions, providing multi-scale image analysis. The fusion of wavelets with Fourier transforms enables a comprehensive examination of both the time and frequency domains.

Traditional methods of image analysis that exclusively leverage Fourier or wavelet transforms have displayed potential in the identification of specific fault types (2). Nevertheless, these techniques grapple with limitations when striving to accurately capture intricate variations or isolated anomalies. This research endeavor introduces a strategy that harnesses the complementary attributes of the wavelet and Fourier methods to address these constraints. Our proposed approach seeks to enhance the representation of image features and augment the distinction between authentic product variations and potential defects. It does so by uniting the multi-resolution analysis capabilities of the wavelet transform with the frequency domain analysis prowess of the Fourier transform. Through this integrated methodology, our study endeavors to significantly bolster the fault detection capabilities of image analysis systems employed in manufacturing quality control.

This transformation has been widely used in image processing and pattern classification, as demonstrated by Zahn (3) in their analysis of closed curves of aircraft using Fourier descriptors based on similarity between curves. The wavelet transform can improve the Fourier transform by enabling simultaneous analysis of signal information in both time and frequency domains; thus, enhancing the ability to capture local signal characteristics quickly. Chen (4) utilized a polar coordinate pattern (h, r) with the center of pattern mass as the origin. They then applied the Fourier transform along the polar angle axis h, and the wavelet transform along the r axis, resulting in extracted features that are invariant to translation, rotation, and scaling. This method was applied to a database of 85 items, and the results showed that the Fourier-Wavelet representation provided better information. Liu (5) proposed a alternative machine vision approach for quantitatively estimating and monitoring the appearance and aesthetics of manufactured products. The approach includes three main steps: (a) wavelet-based textural feature extraction from product images, (b) estimation of measures of product appearance through subspace projection of texture features, and (c) monitoring of the appearance in the latent variable subspace of textural features. Yadav et al. (6) conducted experiments on retrieval and classification of image-based objects using three techniques: Fourier Descriptor (FD), General Fourier Descriptor (GFD), and Fourier-Wavelet Descriptor (WFD). Classification results were compared among all three techniques, and it was found that WFD performed better than FD and GFD techniques The authors propose a method that integrates wavelet analysis and Fourier analysis to enhance the detection and diagnosis of quality issues in manufacturing systems. The technique is applied to a case study in the automotive industry, demonstrating its effectiveness in identifying and resolving quality problems. Hosntalab et al. (7) propose a multi-stage technique to classify teeth in multi-slice CT (MSCT) images. The proposed algorithm consists of the following three stages: segmentation, feature extraction and classification. The employed WFD approach was successful to discriminate and numbering of the teeth in the presence of missing teeth. Münch et al. (8) presented a fast, powerful and stable filter based on combined wavelet and Fourier analysis for the elimination of horizontal or vertical stripes in images and compared with other types of destriping filters. Strict separation between artifacts and original features allowing both, suppression of the unwanted structures and high degree of preservation of the original image information is endeavoured. The results are validated by visual assessments, as well as by quantitative estimation of the image energy loss. Ward et al. (9) discusses the importance of maintaining the accuracy and completeness of in-formation conveyed through visual analytics, as well as the impact of data filtering, sampling, and other transformations on the decision-making process. Megahed et al. (10) propose a new spatiotemporal method for monitoring image data that considers spatial and temporal information. According to the type and purpose of image monitoring in this research, which is to monitor the tile production process, two types of image monitoring conditions have been used: uniformity and comparison with a specific pattern. Chen et al. (11) proposed a method for EEG seizure detection by decomposing EEG signals for up to six wavelet scales without downsampling. Experiments demonstrate that the proposed method is comparable and sometimes better than our previous dual-tree complex wavelet- Fourier method for the University of Bonn EEG database. Knešaurek (12) proposed an improved brain image capture technique using a combined Fourier-wavelet approach. The study results showed that the quality and quantification of 3D brain images can be significantly improved by using the Fourier-wavelet approach com-pared to the original images. Xia et al. (13) compares the effectiveness of wavelet and Fourier transforms for feature extraction in image analysis. The results show that the Wavelet-Fourier method outperforms the individual transforms, demonstrating its potential for enhancing image analysis tasks. Al-Salman et al. (14) proposed a statistical model and Fourier-wavelet analysis for detecting dukes. Ten properties were extracted and used in a method to identify sleep spindles, achieving accuracy, sensitivity, and diagnosis rates of 97.9%, 98.98%, and 97.8% with a window size of 0.5 seconds. You et al. (15) proposed a combined Fourier-Wavelet approach for dynamic analytical solutions. Wavelet transform effectively solves the inverse Fourier transform in physical problem-solving, benefiting flexible multilayer structures and risk management in civil engineering. According to Biglari et al. (16), a combined feature selection technique significantly reduces the number of features with only a slight decrease in prediction accuracy. Fattahzadeh and Saghaei (17) explored the application of video and satellite sensors for process monitoring, emphasizing the need for sequential images to detect complex process changes. They proposed a statistical approach using RIDGE regression and a Q control chart to monitor processes with temporal correlations and high-dimensional data, applicable to both color and gray images. Experimental results, compared to existing methods, demonstrate the method's effectiveness in detecting process changes. Santosh and Barpanda (18) developed an alternative method for glaucoma classification using wavelet-based PCA and enhanced retinal images. The method achieved highly satisfactory results in glaucoma diagnosis and classification. Abdul-Kareem and Al-Jawher (19) developed a new medical image encoding technique by combining discrete wavelet transform, fast Fourier transform, Arnold transform and two multidimensional chaotic systems. The proposed algorithm utilized two real-value transformations: the Discrete Fourier Transform (DFT), which generates real and imaginary coefficients, and the Discrete Wavelet Transform (DWT), which produces four sub-bands, which posed a significant challenge in retrieving images following the encoding and decoding procedure.

The literature review has illuminated the substantial contributions of the Wavelet-Fourier transform. As we move forward in this research, it is evident that the Wavelet-Fourier approach stands as a powerful tool for feature extraction, particularly in the context of image-based quality monitoring and statistical process control.

The research in this field has shown remarkable advancements in the detection and characterization of defects, anomalies, and patterns in data, leading to enhanced performance, reliability, and efficiency in numerous applications. One area of society that greatly benefits from the combined wavelet-Fourier technique is manufacturing. By utilizing this approach, companies are able to detect and diagnose faults or quality issues in their production processes more efficiently and accurately. This helps minimize waste, reduce costs, and improve overall product quality, ultimately leading to customer satisfaction and a competitive edge in the market. Additionally, further investigation into the optimization of the combined wavelet-Fourier technique for real-time applications would be valuable. As technology advances, there is an increasing demand for efficient and rapid data analysis.

Our research primarily aims to address a significant challenge in industrial quality control: the improvement of fault detection in image analysis. To tackle this challenge, we introduce an innovative combined technique, the Wavelet-Fourier method, designed to overcome the limitations associated with individual methods. This novel approach shows great promise in enhancing the identification and classification of defects, ultimately leading to enhanced efficiency and precision within quality control processes in manufacturing industries. By extracting valuable information from images, we empower the use of this data in control charts, allowing proactive measures to prevent and resolve process-related issues. Our study applies the Wavelet-Fourier method to analyze tile images acquired from the production line, using the results for informed decision-making during production and real-time monitoring of the production line. Subsequent sections delve into the feature extraction methods employed and comprehensively discuss the advantages inherent to the Wavelet-Fourier approach. Our research findings, obtained through simulation, robustly affirm the efficacy of our proposed method, illustrating its potential as a valuable tool for image analysis and quality control spanning diverse industries. The subsequent sections describe the feature extraction methods utilized and comprehensively discuss the advantages associated with the Wavelet-Fourier approach. The research findings, obtained through simulation, further substantiate the efficacy of the proposed method, demonstrating its potential as a valuable tool for image analysis and quality control across diverse industries.

2. EXTRACTING THE FEATURES OF IMAGES

Effective use of high-volume image data can significantly improve the performance of quality monitoring methods in manufacturing and service systems. In addition, the high-density information on this type of data can provide more practical guidance in identifying the root cause of process changes and thus reduce the time it takes to return to a controlled state.

Specific features of each image are analyzed to identify differences and similarities. In digital images, pixels represent numeric values that vary based on color (black or gray). The arrangement of pixels creates a matrix with varying numbers, equal to the number of pixels in the image. Numerical calculations can be performed on pixel values, resulting in distinctive features. These features enable pattern recognition, texture recognition, and estimation of size and quantity.

An image is typically represented as a function f (x, y), where x and y are the spatial coordinates of the image, and the value at each location (x, y) is referred to as the intensity or brightness. In digital images, the intensity values are typically non-negative integers. For example, in gray images, f (x, y) can take any integer value between 0 (representing black) and 255 (representing white) for 8-bit images, or between 0 and 65,535 for 16-bit images. It should be noted that as the resolution of the image increases, the number of pixels (i.e., x, y pairs) in the image also increases, which affects the computational complexity.

This results in a much higher pixel intensity range, which can be computationally infeasible for certain applications. To address this, preset subsets called Regions of Interest (ROIs) are often used. ROIs are square-shaped regions within the image where only the mean intensities of the pixels in those regions are used as the basis for calculations, rather than examining every single pixel. This reduces the computational complexity while still providing useful information for analysis. Therefore, the kth image can be represented as a p-variate vector ((p) represents the number of windows (pROI)) based on the mean intensities of the pixels in the ROIs (Equation 1):

$$\bar{X}_{(k)} = \left(\bar{X}_{(k)}(1), \bar{X}_{(k)}(2), \dots, \bar{X}_{(k)}(p)\right) \tag{1}$$

In Phase I, the process is assumed to be in control, following a known or accurately estimated normal distribution with mean $\mu_{0.K}$ and variance ${\sigma^2}_K$. Phase II's purpose is to vigilantly monitor the process, aiming to detect specific causes that might induce a shift in the process output mean.

Control charts, encompassing Shewhart control charts and more advanced statistical process control (SPC) charts, play a pivotal role during Phase II. These charts rely on statistical methods and historical data to assess whether the current process operates within the expected range. If a specific cause leads to a shift in the process output mean, it can lead to deviations from the anticipated distribution $N(\mu_{0.K}.{\sigma^2}_K)$. The control chart promptly identifies such deviations as out-of-control signals, signaling the need for investigation and corrective actions to restore process stability. Employing control charts in Phase II is fundamental for effective process monitoring and quality control, enabling the timely mitigation of process variations to maintain quality standards.

The maximum likelihood estimates $(\hat{\mu}_1)$ and $(\hat{\tau})$ are obtained by maximizing the likelihood function over the parameter space, considering the past m images. The window of the past m images is used to capture the historical information and assess the likelihood of a mean shift based on the observed data. The generalized likelihood ratio statistic is then compared to a predefined threshold or control limits to determine if a mean shift has occurred. This statistical measure is compared to a predefined threshold or control limits to determine if a mean shift has occurred. If the generalized likelihood ratio statistic exceeds the threshold, it indicates evidence of a mean shift, and further investigation and corrective actions may be required to bring the process back under control (20):

$$R_{k,m} = \max_{0\le\tau<k} \frac{k-\hat{\tau}_k}{2\sigma_0^2}\left(\hat{\mu}_{1,\hat{\tau}_k,k} - \mu_0\right) \tag{2}$$

It is worth mentioning that the choice of threshold or control limits for the generalized likelihood ratio statistic depends on the desired level of sensitivity and specificity in detecting mean shifts, and it may require careful consideration and validation based on the specific process and application. Additionally, the selection of the window size (m) for considering the past images also requires careful consideration to capture relevant historical information without including too much noise or irrelevant data. It is important to note that the specific values and parameters used in Equation 2, such as $k, \tau, \sigma_0, \hat{\mu}_{1.}\hat{\tau}_k, \mu_0$, would depend on the particular context and statistical analysis conducted in the study.

In this research, image transformations were selected to extract properties from images. Image transformations involve transforming the revealed objects or the entire image to describe their appearance. The image features are then introduced using all or part of the transform coefficients.

While traditional methods often focus on the argument space (time or place) to extract information from signals or images, it is not always the most effective approach. In many cases, valuable information about signals or images lies in their frequency range. Mathematical transformations applied to signals or images can help to obtain this frequency-based information that may not be easily retrievable from the primary signal or image in the argument space.

The wavelet transform is a crucial tool for transforming signals or images from the argument space to the frequency space. To grasp the concept of the wavelet transform, it is essential to first comprehend the Fourier series. The Fourier series represents periodic signals or functions as a sum of various sinusoidal components with distinct frequencies and amplitudes. It is widely employed to analyze signal or image frequency content in the frequency domain.

2. 1. Frequency Domain and Fourier Transform

Fourier series is a mathematical representation that expresses any periodic function as a sum of simple oscillating functions, such as sines, cosines, or mixed exponential functions. By extending a function as a Fourier series, its frequency components can be obtained, allowing for analysis of its frequency content (21).

In engineering and signal processing, functions are often defined based on time and space, as they represent various physical phenomena. Fourier transform is widely used in various areas of physics, including electronics, electromagnetics, wave physics, and geoscience, such as geophysics (22).

In Fourier transform, frequency and time information cannot be simultaneously obtained. To address this limitation, the Short-Time Fourier Transform (STFT) can be used. In the STFT, the signal's short time interval is divided into small segments, where the signal is assumed to be stationary within each segment. A window function (w) is applied to each segment, which determines the effective domain of the window function and is typically chosen to be equal to the length of the segment (23).

2. 2. Area of Location-scale and Wavelet Transform

Wavelet theory, a modern addition to applied mathematics, has found widespread applications across diverse fields like seismology, signal processing, telecommunications, image processing, computer vision, genetics, medicine, criminology, and telecommunications. Wavelets, a set of mathematical functions, are instrumental in dissecting continuous signals into their frequency components, where each component's detail corresponds to its scale. This analysis leverages the wavelet transform, which breaks down a function using translated and scaled versions of a finite-length function known as the mother wavelet. A key strength of wavelet analysis is its multi-resolution nature, allowing it to extract diverse information from data at varying scales, both in the location and frequency domains simultaneously (24). The fundamental concept behind using wavelet transform for image analysis involves breaking down the image into multiple levels of detail. This process entails dividing the image into distinct domains and generating various sub-images representing different frequency components (25).

They have been used in fields such as genetics and medicine for analyzing DNA sequences, in criminology for pattern recognition in forensic analysis, and in telecommunications for signal compression and noise reduction (26). Wavelet transform $w_x(a,b)$of a continuous signal x (t) is defined in Equation 3:

$$w_x(a,b) = \frac{1}{\sqrt{a}}\int_{-\infty}^{+\infty} x(t)\psi^*(\frac{t-b}{a})dt = \left(x(t).\psi_{a.b}(t)\right) \qquad (a,b)\epsilon R^2 \tag{3}$$

In wavelet theory, the wavelet transform is obtained by multiplying the primary signal, denoted as x(t), with a translated and scaled version of a function called the mother wavelet, denoted as $\psi_{a.b}(t)$,where (a) and (b) are the scaling and translation parameters, respectively. The wavelet transform measures the similarity between the frequency content of the signal and the frequency content of the wavelet function.

The mother wavelet, ψ(t), is a function that is usually chosen to have specific properties, such as being localized in both time and frequency domains. The Haar wavelet, known for its orthogonality and symmetry, is often used for decomposing signals or images (26).

The Haar wavelet is defined in the time domain, with different values within specific intervals and 0 outside. It is piecewise constant and has compact support from 0 to 1, capturing abrupt changes in the signal.

2. 3. Wavelet-Fourier Descriptor (WFD)

The Wavelet-Fourier Descriptor (WFD) is a set of coefficients obtained from Wavelet Transform (WT) of an object's signal. WT is a type of multiple resolution transform that decomposes a signal or image into different levels of detail or resolution (27). In Equation 4 the WT of an image with boundary z(t) is defined (28):

$$\mathrm{WT}(a,b) = \int z(t)\psi^*\left(t-\frac{b}{a}\right)dt \tag{4}$$

where a and b are the parameters of scale and translation, respectively, ψ* is the complex conjugate of the wavelet function ψ, and z(t) is the boundary or contour of the image.

The WFD coefficients are obtained from the WT by extracting the coefficients corresponding to the limit function of the object. These coefficients represent the frequency components of the object's signal at different scales and translations, and can be used as descriptors to characterize the object's shape or features (29). WFD is a powerful tool for image analysis and processing, as it captures the frequency information of an object's signal at different scales and translations, allowing for more detailed and accurate analysis of the object's shape, texture, and other features.

This phrase specifies wavelet coefficients on the boundary of z (t) in scale a and position b. For an optimal scale a, a set of $\mathrm{WT}(\mathrm{a},\mathrm{b})$ that is wave coefficients is obtained. WFDs are selected coefficients that are obtained based on $\mathrm{WT}(\mathrm{a},\mathrm{b})$ wave coefficients and are determined as follows in Equation 5:

$$F^{a}(k)=\frac{1}{N}\sum_{b=0}^{N-1}\mathrm{WT}(\mathrm{a},\mathrm{b})\exp\left({}^{-j2\pi b}/_{N}\right) \qquad (5)$$

It should be noted that the choice of image feature extraction technique depends on the specific requirements of the application, the type of images being analyzed, and the desired invariance properties. Each technique has its advantages and limitations, and performance evaluation should be done based on the specific criteria relevant to the application at hand [30, 31]. According to the studies, we tried to use the same advantages of both Fourier transform and wavelet.

3. SIMULATION STUDIES

The research employs the WFD algorithm for image analysis. It starts with a 250×250 pixel input image, typically a tile from a production line. The WFD algorithm combines wavelet and Fourier techniques, capturing both time and frequency information. It consists of two frequency descriptor levels: low-frequency for general characteristics and high-frequency for finer details.

The image undergoes hierarchical decomposition via recursive wavelet transforms into multiple frequency bands, resulting in a multi-scale image representation. Different bands reveal different levels of detail, enhancing image analysis (30). In this study, the Haar wavelet is specifically used for performing the wavelet transform on the image. The Haar wavelet operates by systematically replacing adjacent pairs of steps in the signal with wider steps, following a specific wavelet methodology (31).

The WFD transformation is a feature extraction technique for tile images. Key steps include:

1. Image Loading: Load a 250×250 pixel input image representing a tile from the production line.
2. Haar Wavelet Transform: Apply the Haar wavelet transform, separating the image into approximation (low-frequency) and detail (fine-detail) coefficients.
3. Fourier Transform of Wavelet Coefficients: Perform Fourier transform on wavelet coefficients, combining multi-scale representation and Fourier analysis.
4. Phase Discarding: Discard the phase part of the Fourier transform for rotation invariance, retaining magnitude information.
5. Normalization: Normalize feature vectors for scale invariance.
6. Low-Frequency and High-Frequency Descriptors: Obtain both low-frequency (general characteristics) and high-frequency (fine details) descriptors.
7. Storage of WFD Coefficients: Collect and store WFD coefficients representing combined wavelet and Fourier information.
8. Comprehensive Set of Coefficients: Repeat the process for all scales and positions to capture various image details.
9. Utilizing WFD Coefficients: Apply the obtained WFD coefficients for tasks like feature extraction, image matching, etc., benefiting from scale, rotation, and translation invariance.
10. Performance Evaluation: Assess the WFD algorithm's effectiveness in image processing tasks, considering specific criteria or metrics, e.g.

WFD transformation on tile images from the production line, it becomes possible to extract meaningful features and analyze the characteristics of the tiles in a comprehensive and invariant manner. Figure 1 represents the algorithm of the WFD technique on the input image.

The WFD algorithm applies approximation and detail coefficients recursively, capturing low-frequency components and high-frequency components with finer details (32). Figure 2 demonstrates the effectiveness of the wavelet transform, showcasing the different frequency bands obtained.

In wavelet transform, image decomposition produces distinct subbands: lowpass approximation, horizontal detail, vertical detail, and diagonal detail. The lowpass approximation represents the image's general

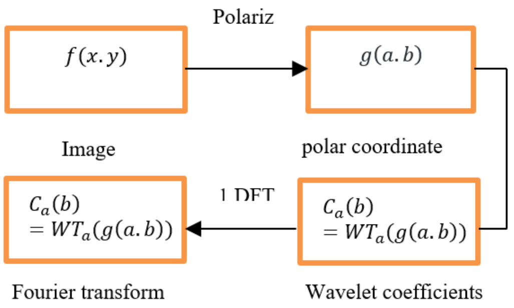


Figure 1. WFD Technique Algorithm

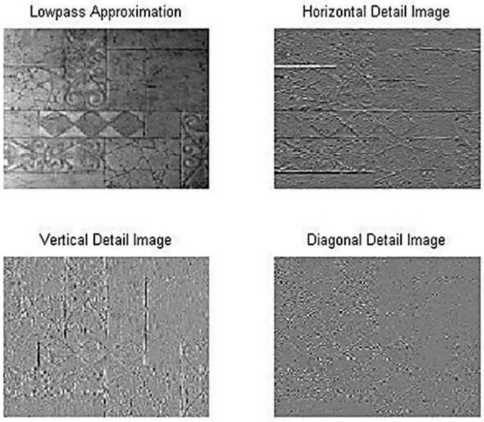


Figure 2. WFD transformation on tile image

characteristics and coarse-scale variations. The detail subbands focus on specific orientations and capture finer variations, including lines, edges, and textures. These subbands enable the analysis of the image's frequency content and spatial information at different scales (33).

This multi-scale representation is a significant advantage of using the wavelet transform in WFD algorithm.

4. CASE STUDY

First, an image of a tile obtained using an Machine Vision System (MVS) is prepared. In the next step, the received image is converted to a grayscale image. Then, we generate the actual image of the tile by subtracting the background intensity from the image. To reduce calculations, the image is resized to 250×250. Subsequently, a similar image is produced by adding noise to the image. A random noise image was added to each pixel. The random noise added to each pixel follows a Poisson distribution with a mean pixel value. In this study, the performance of the combined Wavelet-Fourier technique was thoroughly evaluated using three key criteria: average run length, standard deviation of run length, and accuracy of estimation of change point. Additionally, the size and location of faults were taken into account. The corresponding numerical results can be found in Tables 1 and 2, which provide comprehensive information regarding the performance of the technique.

To conduct the performance evaluation, the technique was tested on a total of 150 experimental fault scenarios. These scenarios encompassed a range of variables, including three fault centers, five fault sizes, and ten intensity shifts (Δ). By incorporating such diverse factors, the study aimed to thoroughly assess the effectiveness and robustness of the Wavelet-Fourier technique.

To ensure a fair comparison, the settings utilized in a previous work conducted by Megahed et al. (10) were adopted as the baseline for comparison. This allowed for a meaningful evaluation of the improvements and advancements achieved in this study.

In the pursuit of advancing our image-based quality control and monitoring methodology, a pivotal aspect of our research involves the systematic generation of controlled errors within our image dataset. These errors are meticulously designed to simulate real-world defects and anomalies in tiles, enabling us to rigorously assess the performance and robustness of our quality monitoring techniques. Figure 3, demonstrates detailed visual representation of our error generation and application process we designated in our work. The sequence of operations commences with the preparation of the image, followed by the precise creation of errors with predetermined characteristics. Subsequently, these errors are intentionally introduced into the image dataset, providing us with a controlled testbed for evaluating the efficacy of our quality control algorithms.

TABLE 1. Simulation results for 15 × 15 fault size

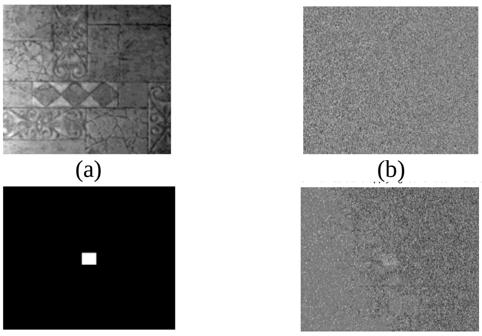
15*15										
$(\varepsilon_\tau) > 2$	$(\varepsilon_\tau) < -2$	$0 < (\varepsilon_\tau) \leq 2$	$(\varepsilon_\tau) = 0$	Std (ε_τ)	Med (ε_τ)	mean (ε_τ)	MRL	ARL	Δ	Fault Center
0	0	1	99	0.10	0	-0.01	1	1.11	-10	(125*125)
0	3	28	69	0.87	0	-0.17	3	3.09	-5	
9	6	36	49	1.66	0	0.17	6	6.47	-3	
54	7	25	14	12.11	3	8.35	11.50	16.18	-2	
97	1	1	1	89.79	78.50	102.96	84.50	108.97	-1	
96	1	2	1	136.50	116	149.75	123.5	154.75	1	
83	4	9	4	25.03	18	23.930	25.50	31.41	2	
24	4	32	40	3.92	0	1.9000	8	8.81	3	
2	1	25	72	1.09	0	-10.0	3	3.57	5	
0	0	0	100	0	0	0	1	1.17	10	

0	0	3	97	0.17	0	-0.03	1	1.10	-10	(188*206)
0	2	21	77	0.86	0	-0.15	3	2.87	-5	
6	7	43	44	2.43	0	0.23	6	6.52	-3	
59	2	25	14	13.29	4	8.98	12	16.38	-2	
98	0	1	1	118.24	88	122.80	91	127.92	-1	
97	0	3	0	99.57	79.50	105.02	84.50	110.68	1	
64	4	16	16	10.26	4	8.2	11	15.08	2	
19	0	33	48	1.96	0	0.90	7	6.66	3	
0	3	16	81	0.72	0	-0.10	3	2.67	5	
0	0	0	100	0	0	0	1	1.07	10	
0	1	0	99	0.80	0	-0.08	1	1.01	-10	(158*78)
0	4	13	83	1.10	0	-0.26	2	2.24	-5	
3	5	34	58	1.41	0	0.17	5	4.95	-3	
46	2	29	23	7.29	2	4.87	9	11.75	-2	
92	3	3	2	63.32	47	67.91	54.50	74.14	-1	
94	1	4	1	79.46	61	82.72	68	88.80	1	
43	4	26	27	7.06	2	4.20	10	11.46	2	
4	3	43	50	1.38	0	0.03	5	4.87	3	
0	4	10	86	1.29	0	-0.25	2	2.23	5	
0	0	0	100	0	0	0	1	1	10	

TABLE 2. Simulation results for 20 × 20 fault size

20*20										
$(\varepsilon_\tau) > 2$	$(\varepsilon_\tau) < -2$	$0 < (\varepsilon_\tau) \leq 2$	$(\varepsilon_\tau) = 0$	**Std** (ε_τ)	**Med** (ε_τ)	**mean** (ε_τ)	**MRL**	**ARL**	**Δ**	**Fault Center**
0	0	0	100	0	0	0	0	1	-10	(125*125)
0	1	6	93	0.482	0	-0.10	2	1.56	-5	
2	2	26	70	1.147	0	-0.24	4	3.71	-3	
12	6	37	45	2.160	0	0.32	6	6.42	-2	
82	2	10	6	32.20	19	30.03	26.5	36.85	-1	
98	1	0	1	63.30	51	67.78	57	73.57	1	
25	2	31	42	3.040	0	1.49	7	7.99	2	
3	3	32	62	1.132	0	-0.030	4	3.65	3	
0	0	6	94	0.301	0	-0.010	2	1.72	5	
0	0	0	100	0	0	0	1	1	10	
0	0	0	100	0	0	0	1	1	-10	(188*206)
0	1	4	95	0.748	0	-0.080	1	1.38	-5	
1	0	18	81	0.722	0	0.060	3	3.01	-3	
9	2	37	52	2.066	0	0.340	6	5.91	-2	
77	4	10	9	39.73	17	28.24	25	35.49	-1	
85	3	6	6	32.39	18.50	26.53	24	33.79	1	
9	2	45	44	1.664	0	0.40	5	5.62	2	
0	4	24	72	1.073	0	-0.20	3	2.81	3	
0	1	5	94	0.634	0	-0.11	1	1.43	5	
0	0	0	100	0	0	0	1	1	10	

0	0	0	100	0	0	0	1	1	-10	(158*78)
0	0	2	98	0.142	0	0	1	1.15	-5	
0	1	18	81	0.595	0	-0.10	2	2.38	-3	
2	5	24	69	1.351	0	-0.18	5	4.70	-2	
79	3	11	7	17.35	12	16.90	20	24.47	-1	
77	2	12	9	16.85	9.50	15.11	18.5	22.65	1	
2	4	33	61	1.459	0	-0.070	5	4.83	2	
0	5	6	89	1.181	0	-0.33	2	2.51	3	


(a) (b)

(c) (d)

Figure 3. (a) Image, (b) Processed image, (c) Creating fault, (d) Image with fault

The study employed a sophisticated ROI production heuristic algorithm, which consisted of three distinct stages. In the first stage, grid points were strategically specified on the image based on the ratio between the number of pixels and the desired number of grid points along both the width and length dimensions. The specified ratio ranged from 25 to 100. As an example, for a 250×250 image, a ratio of 25 was selected for each direction, resulting in a total of 100 grid points evenly distributed across the entire image.

The research involved a multi-stage process to evaluate the performance of the combined Wavelet-Fourier technique in detecting faults in experimental images. In the second stage, the smallest size for ROI was determined, factoring in camera resolution, zoom level, camera position relative to the object, and image processing compression. This ensured even the smallest faults were adequately analyzed. ROIs were generated using a specific methodology, starting with the smallest size and incrementally increasing. The process continued until the image boundary was reached, following spatial guidance. The results from these ROIs were rigorously analyzed using various criteria, including average run length, standard deviation of run length, accuracy of change point estimation, and fault size and location. These criteria provided quantitative insights into the technique's effectiveness in fault detection and analysis, enabling a comprehensive evaluation of its performance.

In the second phase of the study, the proposed technique underwent rigorous evaluation using MATLAB software, primarily focusing on calculating Average Run Length (ARL) values as a performance measure. To ensure a robust assessment, the Median Run Length (MRL) criterion was also employed, offering greater reliability by being less sensitive to outliers.

To establish a reference point for the in-control condition, the ARL value was adjusted to 200, serving as a baseline for the average number of observations required to detect system faults or changes. In this evaluation, a simulation program with 1000 runs determined an MRL value of 145. This result indicates that, on average, the proposed technique detected faults at a shorter run length compared to the in-control condition, showcasing its effectiveness in prompt fault identification.

The evaluation further included comprehensive computer simulations, considering various parameters like size, quantity, and location of intensity shifts, to emulate practical fault detection scenarios. Over 1000 repetitions were performed for each parameter combination to ensure result reliability. To maintain evaluation integrity, steps were taken to mitigate environmental factors that could impact image quality, ensuring the assessment provided accurate and unbiased results.

During the comparative analysis in phase 2, the MRLvalues obtained from the simulation study were utilized. The analysis focused on assessing the proposed technique's ability to accurately detect faults when faced with different scenarios involving variations in fault size, quantity, and location of intensity shifts.

By thoroughly examining the results obtained from the simulation study, a comprehensive understanding of the proposed technique's performance was achieved. The analysis provided insights into the technique's effectiveness in detecting faults across a range of conditions, allowing researchers and readers to gain valuable knowledge about its capabilities and limitations. This rigorous evaluation process ensured that the proposed technique's performance was thoroughly assessed and its ability to detect faults under diverse conditions was accurately characterized.In this study, several factors were considered to conduct a comprehensive case study analysis. The size of the smallest ROI was determined to be 22×22 pixels, providing a baseline for evaluating the proposed method's performance. Three defect centers were selected at coordinates (188×206), (78×158), and (125×125), representing different positions within the image. These defect centers were analyzed to assess the effectiveness of the proposed method in detecting faults. Upon analyzing the results for the mentioned defect centers, it was observed that the proposed method demonstrated superior performance for the centers at (188×206) and (78×158). However, the results for the center at (125×125) were comparable to previous studies, indicating that the proposed method can yield similar outcomes in such cases. This finding suggests that employing combined methods of image extraction allows for a more comprehensive investigation of different areas within the image. Additionally, statistics such as the change point ($\varepsilon_{\tau} = \hat{\tau} - \tau$) were reported in the Table. Change point analysis is a statistical technique used in process monitoring to identify significant shifts or changes in a process's behavior. It helps detect sudden variations in process parameters or quality measurements, highlighting potential alterations in process dynamics or operating conditions. Change point analysis is a valuable tool for monitoring processes, enabling timely interventions, quality improvements, and waste reduction. However, it requires careful consideration of data quality, appropriate statistical methods, and interpretation of results in the context of the specific process being monitored. For example, in the case of a fault size of 15×15 at the center of a 206×188 image, with a $\Delta = -3$, the average run length is 6. Conversely, for a fault size of 20×20 with the same specifications, it is reduced to 3. A noteworthy point to consider is that both of these defect sizes are smaller than the smallest window. Additionally, it can be claimed that larger defect sizes are detected more quickly.

In accordance with Megahed et al.'s (10) work, the proposed change point technique demonstrated generally robust performance across various fault sizes. However, similar studies have indicated a notable decrease in performance when dealing with minor intensity shifts and fault sizes of 15×15 pixels or smaller. This aligns with expectations, as tiny intensity shifts can be challenging to precisely diagnose, particularly when barely perceptible to the human eye. The study revealed that the proposed method excelled in detecting faults within the range of intensity shifts from -10 to 10, irrespective of the defect size. This highlights the method's effectiveness in identifying faults within this specific intensity shift range. Table 3 presents a comparative analysis between this study's results and Megahed et al.'s (10) work, with a focus on a 15×15 pixel fault centered at (125×125). The findings indicate that the proposed method performs notably better for faults situated at the sides or edges of the image or ROI. This underscores the approach's particular efficacy in detecting faults in these peripheral areas.

An essential observation from the analysis pertains to the relationship between the size of the smallest ROI and the minimum fault size requiring identification. Properly sizing and positioning the ROI in relation to the fault size are critical considerations that significantly influence detection accuracy and reliability. Overlapping between the ROI and the fault can introduce complexities and potentially compromise detection accuracy. This emphasizes the need for meticulous attention to ROI design and positioning to ensure precise fault detection. Addressing the potential impact of overlapping factors on detection accuracy is crucial for enhancing fault detection techniques' effectiveness and overall accuracy in fault identification. The comparison in Table 3 not only highlights the superior performance of the proposed method for peripheral faults but also underscores the critical importance of appropriately sizing and positioning the ROI to mitigate overlapping factors that may affect detection accuracy. These insights provide valuable guidance for future research in fault detection, emphasizing the need for meticulous ROI design and positioning consideration.

The Dice Similarity Coefficient (DSC) calculates the spatial overlap between the target areas A and B. It provides a numerical value ranging from 0 to 1, where a value of 1 indicates perfect agreement or complete overlap between the target areas, while a value of 0 represents no overlap or similarity (34). The DSC is calculated using Equation 6 mentioned in the study. This information is crucial in determining the reliability and effectiveness of the method for fault localization and provides insights into the potential applications and performance of the proposed technique (34).

$$\mathrm{DSC}\,(\mathrm{A}\,,\mathrm{B}\,) = \frac{(2\times|A\cap B|)}{|A|+|B|} \qquad (6)$$

In the realm of fault detection and localization, the DSC serves as a crucial evaluation metric for assessing the accuracy and congruence between estimated and reference fault locations. A higher DSC value signifies more precise fault estimation, while a lower value implies reduced accuracy. To enhance the assessment, two additional metrics, the Median of Dice Similarity Coefficient (MDSC) and the Maximum Achievable Dice Similarity Coefficient (MADSC), are employed. MDSC offers a benchmark for gauging typical similarity between estimated and actual fault areas, independent of heuristic techniques for generating ROI. Meanwhile, MADSC quantifies the upper limit of attainable similarity, acting as a reference point, especially for specific fault sizes. The incorporation of MDSC and MADSC enables a comprehensive evaluation of the method's precision in fault size and location estimation, shedding light on its potential and limitations within the given conditions. In essence, these measures enrich the assessment of the proposed method's fault estimation performance.

For example, for faults of size 20×20, MADSC is calculated as follows:

$$\frac{(2*|20*20|)}{|20*20|+|(22*22)|} = 0.905$$

Table 4 showcases the performance of the method in estimating the time and size of faults under intensity shifts. To evaluate the method's performance, the ratio of $\left(\frac{MDSC}{MADSC}\right)$ is employed as a measure. This ratio serves as an indicator of how well the method performs compared to the maximum achievable similarity between the estimated and actual fault areas. It is important to note that this assessment assumes alignment between the ROI center and the fault of the MADSC benchmark structure.

The ratio is calculated for all three fault centers considered in the study. For the two side centers, where the intensity shifts are high, the ratio is predominantly above 0.5. This suggests that the method performs relatively well in these scenarios, exhibiting satisfactory performance in estimating both the time and size of faults.

Similarly, for the center located at (125×125) with faults larger than the smallest ROI, the ratio is higher than 0.5, indicating a favorable performance. However, it is worth noting that cases where |Δ|=1 and the fault size is 10x10 might result in an MDSC value of 0. This is expected because the MRL value in these cases is very close to the MRL under control conditions. Consequently, most of these signals are likely to be false alerts attributed to the shift, leading to a lower coverage benchmark.

In summary, Table 4 provides insights into the method's performance in estimating the time and size of faults under intensity shifts. The ratio of MDSC to MADSC serves as a comparative measure, showcasing the method's performance relative to the maximum achievable similarity. The observed results indicate favorable performance for faults located at the side centers and for larger faults at the center (125×125). However, it is important to consider the specific conditions and limitations, as cases with minimal intensity shifts and smaller fault sizes may yield lower MDSC values due to potential false alerts.

5. MANAGEMENT INSIGHT

In the contemporary data-driven landscape, extracting insights from abundant image data is pivotal for manufacturing and service systems. This insight offers a practical roadmap for applying the proposed method:

TABLE 3. Comparison with other research for size faults 15*15

Proposed method								**15*15**
$(\varepsilon_\tau) > 2$	$(\varepsilon_\tau) < -2$	$0 < (\varepsilon_\tau) \le 2$	$(\varepsilon_\tau) = 0$	**Std** (ε_τ)	**Med** (ε_τ)	**mean** (ε_τ)	**Δ**	**Fault Center**
0	3	28	69	0.87	0	-0.17	-5	(125*125)
9	6	36	49	1.66	0	0.17	-3	
54	7	25	14	12.11	3	8.35	-2	
97	1	1	1	89.79	78.50	102.96	-1	
96	1	2	1	136.50	116	149.75	1	
83	4	9	4	25.03	18	23.930	2	
24	4	32	40	3.92	0	1.9000	3	
2	1	25	72	1.09	0	-10.0	5	
Research by Megahed et al. [10]								
0.3	2.2	20.5	77	0	0.851	-0.095	-5	(125*125)
16.6	3.1	36.8	43.5	0	2.623	0.912	-3	
72.3	2.1	15.7	9.9	9	16.577	13.795	-2	
96.9	0.7	1.7	0.7	83	131.78	126.02	-1	
96.6	0.5	2.3	0.6	87.5	123.56	124.44	1	
71.3	1.2	17.5	10	8.5	15.085	12.735	2	
14.6	3.5	38.2	43.7	0	2.461	0.821	3	
0.4	2.4	18	79.2	0	0.875	-0.099	5	

Step 1: Introduction and Context

Highlighting the relevance, this approach underscores the importance of using image data for quality monitoring and process control. It recognizes the need for robust methods in an era of data abundance.

Step 2: Selecting the Most Suitable Method

Emphasizing method selection, the insight advocates the Wavelet-Fourier approach's superiority over individual methods like Fourier or Wavelet transforms. This underscores the critical role of choosing optimal feature extraction methods for effective online process monitoring.

Step 3: Identifying Image Processing Issues and Providing Valuable Information

Underlining the method's capability to identify image processing issues, it emphasizes its ability to estimate change points, fault locations, and sizes. This empowers organizations to promptly address process changes, reducing downtime

Step 4: Integration of Image Data Analysis into Quality Monitoring

Highlighting the potential of integrating image data analysis into quality monitoring, it showcases the profound insights that dense image data can provide, leading to enhanced decision-making, efficiency, cost savings, and customer satisfaction.

Step 5: Conclusion and Overall Benefit

Reiterating the overarching benefits, it underscores how effective image data utilization and advanced feature extraction methods can improve quality monitoring, drive efficiency, reduce costs, and elevate customer satisfaction in today's competitive business landscape.

TABLE 4. Comparison of (MDSC/MADSC) for all 3 fault centers & 10 values intensity shift

	(78*158)		**(188*206)**			**(125*125)**			**center**	
Δ	size	E	Med	$\frac{MDSC}{MADSC}$	E	Med	$\frac{MDSC}{MADSC}$	E	Med	$\boldsymbol{\frac{MDSC}{MADSC}}$
-10	10*10	0.155	0.163	0.477	0.165	0.180	0.524	0.193	0.224	**0.652**
	15*15	0.242	0.243	0.383	0.265	0.270	0.425	0.336	0.368	**0.580**
	20*20	0.334	0.333	0.368	0.371	0.367	0.405	0.424	0.454	**0.501**
	30*30	0.475	0.482	0.482	0.579	0.581	0.581	0.556	0.596	**0.596**
	50*50	0.629	0.632	0.632	0.764	0.765	0.765	0.670	0.658	**0.658**
-5	10*10	0.146	0.149	0.433	0.143	0.163	0.477	0.170	0.224	**0.652**
	15*15	0.230	0.243	0.383	0.259	0.270	0.425	0.308	0.346	**0.545**
	20*20	0.321	0.333	0.368	0.374	0.385	0.426	0.394	0.428	**0.473**
	30*30	0.460	0.467	0.467	0.562	0.581	0.581	0.538	0.553	**0.553**
	50*50	0.624	0.623	0.623	0.755	0.765	0.765	0.685	0.658	**0.658**
-3	10*10	0.122	0.147	0.430	0.114	0.147	0.430	0.140	0.127	**0.370**
	15*15	0.225	0.243	0.383	0.253	0.270	0.425	0.302	0.346	**0.544**
	20*20	0.309	0.333	0.368	0.372	0.389	0.430	0.394	0.428	**0.473**
	30*30	0.450	0.467	0.467	0.537	0.551	0.551	0.499	0.521	**0.521**
	50*50	0.610	0.620	0.620	0.742	0.745	0.745	0.673	0.658	**0.658**
-2	10*10	0.068	0.024	0.071	0.048	0	0	0.084	0.038	**0.111**
	15*15	0.199	0.243	0.383	0.219	0.270	0.425	0.275	0.346	**0.544**
	20*20	0.300	0.333	0.368	0.349	0.389	0.430	0.366	0.403	**0.445**
	30*30	0.427	0.448	0.448	0.518	0.552	0.552	0.494	0.502	**0.502**
	50*50	0.570	0.605	0.605	0.725	0.725	0.725	0.673	0.658	**0.658**
-1	10*10	0.022	0	0	0.009	0	0	0.0149	0	**0**
	15*15	0.102	0.060	0.94	0.089	0	0	0.131	0.101	**0.159**
	20*20	0.242	0.320	0.345	0.241	0.333	0.368	0.297	0.307	**0.339**
	30*30	0.385	0.433	0.433	0.510	0.543	0.543	0.478	0.502	**0.502**
	50*50	0.561	0.588	0.588	0.695	0.732	0.732	0.616	0.652	**0.652**

1	10*10	0.011	0	0	0.0127	0	0	0.0183	0	**0**
	15*15	0.050	0	0	0.113	0.066	0.104	0.158	0.121	**0.191**
	20*20	0.167	0.173	0.191	0.275	0.365	0.404	0.303	0.321	**0.355**
	30*30	0.389	0.443	0.443	0.526	0.559	0.559	0.479	0.585	**0.585**
	50*50	0.567	0.594	0.594	0.694	0.723	0.723	0.632	0.652	**0.652**
2	10*10	0.029	0	0	0.067	0.030	0.086	0.0861	0.063	**0.184**
	15*15	0.198	0.243	0.383	0.206	0.270	0.425	0.275	0.346	**0.544**
	20*20	0.298	0.332	0.367	0.361	0.381	0.422	0.348	0.331	**0.367**
	30*30	0.442	0.456	0.456	0.531	0.551	0.551	0.502	0.540	**0.540**
	50*50	0.582	0.610	0.610	0.723	0.743	0.743	0.665	0.658	**0.658**
3	10*10	0.081	0.089	0.259	0.110	0.147	0.430	0.149	0.134	**0.390**
	15*15	0.223	0.243	0.383	0.247	0.270	0.425	0.306	0.346	**0.544**
	20*20	0.312	0.333	0.368	0.360	0.378	0.418	0.380	0.421	**0.465**
	30*30	0.447	0.467	0.467	0.464	0.575	0.575	0.514	0.540	**0.540**
	50*50	0.608	0.618	0.618	0.748	0.765	0.765	0.682	0.681	**0.681**
5	10*10	0.140	0.147	0.430	0.0154	0.163	0.477	0.164	0.202	**0.588**
	15*15	0.230	0.243	0.383	0.259	0.270	0.425	0.315	0.346	**0.544**
	20*20	0.320	0.333	0.368	0.361	0.378	0.418	0.398	0.428	**0.473**
	30*30	0.462	0.467	0.467	0.563	0.581	0.581	0.535	0.540	**0.540**
	50*50	0.620	0.623	0.623	0.759	0.765	0.765	0.683	0.658	**0.658**
10	10*10	0.154	0.163	0.477	0.157	0.163	0.477	0.195	0.334	**0.652**
	15*15	0.239	0.243	0.383	0.261	0.270	0.425	0.333	0.368	**0.580**
	20*20	0.333	0.33	0.368	0.375	0.367	0.405	0.425	0.454	**0.501**
	30*30	0.475	0.482	0.482	0.577	0.581	0.581	0.567	0.596	**0.596**
	50*50	0.630	0.634	0.634	0.764	0.765	0.765	0.690	0.658	**0.658**

By following these steps, the management insight provides a clear and concise overview of the research findings and their implications for management in manufacturing and service systems.

It highlights the importance of selecting the right method, the benefits of identifying image processing issues, and the value of integrating image data analysis into quality monitoring practices.

6. CONCLUSION

This study has successfully demonstrated the effectiveness of the combined Wavelet-Fourier technique for enhancing fault detection in image analysis for fabrication quality control. The proposed approach not only identifies image processing issues but also provides valuable information for estimating change points, fault locations, and sizes, leading to more efficient resolution and prediction of faults. Through the evaluation of its performance on experimental images, the combined Wavelet-Fourier technique showcased reliable fault detection capabilities, outperforming existing methods, particularly for faults situated at the edges or sides of the image. The comprehensive analysis of evaluation criteria and benchmark measures confirmed the method's superior performance, signifying its potential for diverse applications in manufacturing and quality control. By leveraging image data for statistical process control, manufacturers can make informed decisions to improve overall product quality while achieving cost and time savings in production. The practical implications of this research extend beyond the laboratory, offering a comprehensive solution for fault detection and analysis in various industries. Proper consideration and adjustment of parameters, such as the size and positioning of the region of interest, are crucial to obtaining accurate fault detection results. While the results are promising, it's essential to acknowledge certain limitations and suggest avenues for future research. One limitation of our study is the sensitivity of the proposed method to parameter settings, such as the size and positioning of the ROI. Further investigation into parameter robustness and automated parameter optimization methods could enhance the technique's applicability. The experimental dataset primarily focused on specific fault scenarios. Future research should expand the scope to include a more diverse range of fault types and complexities to validate the method's effectiveness across various industrial applications. Also continued research should aim to refine and optimize the combined Wavelet-Fourier technique. This includes exploring variations of the method and fine-tuning its parameters to make it more adaptable to different image datasets and fault scenarios. Future studies should consider applying the technique to a broader range of image datasets from various industries. This will help assess its versatility and generalizability in different manufacturing and service systems. Investigate the integration of automated parameter tuning and machine learning algorithms to enhance the method's efficiency and accuracy in fault detection, potentially enabling real-time implementation. In summary, while the combined Wavelet-Fourier technique represents a significant advancement in fault detection, addressing its limitations and pursuing further research avenues will be instrumental in harnessing its full potential for the benefit of manufacturing and quality control across diverse industries.

7. REFERENCES

1. Feng J, Fu J, Lin Z, Shang C, Li B. A review of the design methods of complex topology structures for 3D printing. Visual Computing for Industry, Biomedicine, and Art. 2018;1(1):1-16. https://doi.org/10.1186/s42492-018-0004-3
2. Hu G-H, Wang Q-H, Zhang G-H. Unsupervised defect detection in textiles based on Fourier analysis and wavelet shrinkage. Applied optics. 2015;54(10):2963-80. https://doi.org/10.1364/AO.54.002963
3. Zahn CT, Roskies RZ. Fourier descriptors for plane closed curves. IEEE Transactions on computers. 1972;100(3):269-81. https://doi.org/10.1109/TC.1972.5008949
4. Chen G, Bui TD. Invariant Fourier-wavelet descriptor for pattern recognition. Pattern recognition. 1999;32(7):1083-8. https://doi.org/10.1016/S0031-3203(98)00148-4
5. Liu JJ, MacGregor JF. Estimation and monitoring of product aesthetics: application to manufacturing of "engineered stone" countertops. Machine Vision and Applications. 2006;16:374-83. https://doi.org/10.1007/s00138-005-0009-8
6. Yadav RB, Nishchal NK, Gupta AK, Rastogi VK. Retrieval and classification of shape-based objects using Fourier, generic Fourier, and wavelet-Fourier descriptors technique: A comparative study. Optics and Lasers in engineering. 2007;45(6):695-708. https://doi.org/10.1016/j.optlaseng.2006.11.001
7. Hosntalab M, Aghaeizadeh Zoroofi R, Abbaspour Tehrani-Fard A, Shirani G. Classification and numbering of teeth in multi-slice CT images using wavelet-Fourier descriptor. International journal of computer assisted radiology and surgery. 2010;5:237-49. https://doi.org/10.1007/s11548-009-0389-8
8. Münch B, Trtik P, Marone F, Stampanoni M. Stripe and ring artifact removal with combined wavelet—Fourier filtering. Optics express. 2009;17(10):8567-91. https://doi.org/10.1364/OE.17.008567
9. Ward M, Xie Z, Yang D, Rundensteiner E. Quality-aware visual data analysis. Computational Statistics. 2011;26(4):567-84. https://doi.org/10.1007/s00180-010-0226-0
10. Megahed FM, Wells LJ, Camelio JA, Woodall WH. A spatiotemporal method for the monitoring of image data. Quality and Reliability Engineering International. 2012;28(8):967-80. https://doi.org/10.1002/qre.1287
11. Chen G, Xie W, Bui TD, Krzyżak A. Automatic epileptic seizure detection in EEG using nonsubsampled wavelet–fourier features. Journal of Medical and Biological Engineering. 2017;37:123-31. https://doi.org/10.1007/s40846-016-0214-0
12. Knešaurek K. Fourier-wavelet restoration in PET/CT brain studies. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment. 2012;689:29-34. https://doi.org/10.1016/j.nima.2012.06.032
13. Xia Y, Johnson BK, Jiang Y, Fischer N, Xia H. A new method based on artificial neural network, Wavelet Transform and Short Time Fourier Transform for Subsynchronous Resonance detection. International Journal of Electrical Power & Energy Systems. 2018;103:377-83. https://doi.org/10.1016/j.ijepes.2018.06.019
14. Al-Salman W, Li Y, Wen P. Detecting sleep spindles in EEGs using wavelet fourier analysis and statistical features. Biomedical Signal Processing and Control. 2019;48:80-92. https://doi.org/10.1016/j.bspc.2018.10.004
15. You L, Man J, Yan K, Wang D, Li H. Combined Fourier-wavelet transforms for studying dynamic response of anisotropic multi-layered flexible pavement with linear-gradual interlayers. Applied Mathematical Modelling. 2020;81:559-81. https://doi.org/10.1016/j.apm.2020.01.031
16. Biglari M, Mirzaei F, Hassanpour H. Feature selection for small sample sets with high dimensional data using heuristic hybrid approach. International Journal of Engineering, Transactions B: Applications. 2020;33(2):213-20. https://doi.org/10.5829/IJE.2020.33.02B.05
17. Fattahzadeh M, Saghaei A. A statistical method for sequential images–based process monitoring. International Journal of Engineering, Transactions A: Basics. 2020;33(7):1285-92. https://doi.org/10.5829/IJE.2020.33.07A.15
18. Santosh NK, Barpanda SS. Wavelet and PCA-based glaucoma classification through novel methodological enhanced retinal images. Machine Vision and Applications. 2022;33(1):11. https://doi.org/10.1007/s00138-021-01263-w
19. Abdul-Kareem AA, & Al-Jawher, W. A. M. . A Hybrid Domain Medical Image Encryption Scheme Using URUK and WAM Chaotic Maps with Wavelet–Fourier Transform. Journal of Cyber Security and Mobility. 2023:435-64.
20. Reynolds Jr MR, Lou J. An evaluation of a GLR control chart for monitoring the process mean. Journal of quality technology. 2010;42(3):287-310. https://doi.org/10.1080/00224065.2010.11917825
21. Wirsing K. Time frequency analysis of wavelet and Fourier transform. Wavelet theory. 2020.
22. Stein EM, Weiss G. Introduction to Fourier Analysis on Euclidean Spaces (PMS-32). Introduction to Fourier Analysis on Euclidean Spaces (PMS-32).32.
23. Zhang Z, Jing Z, Wang Z, Kuang D. Comparison of Fourier transform, windowed Fourier transform, and wavelet transform

methods for phase calculation at discontinuities in fringe projection profilometry. Optics and Lasers in Engineering. 2012;50(8):1152-60. https://doi.org/10.1016/j.optlaseng.2012.03.004

24. Lin L, Feng L. Comparative analysis of image denoising methods based on wavelet transform and threshold functions. International Journal of Engineering. 2017;30(2):199-206. https://doi.org/10.5829/idosi.ije.2017.30.02b.06
25. Kunttu I, Lepistö L, Rauhamaa J, Visa A. Multiscale Fourier descriptors for defect image retrieval. Pattern Recognition Letters. 2006;27(2):123-32. https://doi.org/10.1016/j.patrec.2005.08.022
26. Mallat S. A wavelet tour of signal processing: Elsevier; 1999.
27. Shen D IH. Discriminative wavelet shape descriptors for recognition of 2-D patterns. Pattern recognition. 1999;32(2):151-65.
28. Illanes A, Esmaeili N, Poudel P, Balakrishnan S, Friebe M. Parametrical modelling for texture characterization—A novel approach applied to ultrasound thyroid segmentation. PloS one. 2019;14(1):e0211215. https://doi.org/10.1371/journal.pone.0211215
29. Matuszewski DJ, Hast A, Wählby C, Sintorn I-M. A short feature vector for image matching: The Log-Polar Magnitude feature descriptor. Plos one. 2017;12(11):e0188496. https://doi.org/10.1371/journal.pone.0188496
30. Bishop CM, Nasrabadi NM. Pattern recognition and machine learning: Springer; 2006.
31. Rao RM. Wavelet transforms: Introduction to theory and applications: Pearson Education India; 1998.
32. Haar A. Zur theorie der orthogonalen funktionensysteme: Georg-August-Universitat, Gottingen.; 1909.
33. Kumar S, Bhandari AK, Raj A, Swaraj K. Triple clipped histogram-based medical image enhancement using spatial frequency. IEEE Transactions on NanoBioscience. 2021;20(3):278-86. https://doi.org/10.1109/TNB.2021.3064077
34. Wang J, Cui L, Xu Y. Quantitative and localization fault diagnosis method of rolling bearing based on quantitative mapping model. Entropy. 2018;20(7):510. https://doi.org/10.3390/e20070510

Persian Abstract

چکیده

این مطالعه بر استفاده از داده‌های تصویری برای کنترل فرآیند آماری و بهبود نظارت بر کیفیت در سیستم‌های تولید و خدمات تمرکز دارد. اثربخشی روش‌های استخراج ویژگی فردی و ترکیبی با رویکرد موجک- فوریه به عنوان مناسب‌ترین روش مورد ارزیابی قرار می‌گیرد. روش پیشنهادی نه تنها مسائل مربوط به پردازش تصویر را شناسایی می‌کند، بلکه اطلاعات ارزشمندی را برای تخمین نقاط تغییر، مکان‌های خطا و اندازه خطا ارائه می‌دهد. این امکان حل و پیش بینی خطا را فراهم می کند و منجر به صرفه جویی در هزینه و زمان در تولید می شود. برای ارزیابی عملکرد روش، تصویری از خط تولید کاشی تحت تبدیل موجک قرار می‌گیرد و به دنبال آن تبدیل فوریه روی ضرایب به‌دست‌آمده انجام می شود. نتایج نشان دهنده برتری روش موجک- فوریه بر روش های فردی مانند تبدیل فوریه و تبدیل موجک است. روش پیشنهادی، عملکرد قابل مقایسه یا بهبود یافته ای را در تشخیص و مکان خطا در مقایسه با تحقیقات مشابه نشان می دهد. این مطالعه پتانسیل استفاده از داده های تصویری برای کنترل فرآیند آماری و نظارت بر کیفیت را برجسته می کند و راه حلی جامع برای تشخیص و تجزیه و تحلیل خطا ارائه می دهد. این یافته ها به پیشرفت در تکنیک های پردازش تصویر کمک می کند و پیامدهای عملی برای افزایش نظارت بر کیفیت در صنایع مختلف دارد. با استفاده از داده‌های تصویر، تولیدکنندگان می‌توانند تصمیمات آگاهانه بگیرند، عملکرد فرآیند را افزایش دهند و کیفیت کلی محصول را بهبود بخشند.

IJE TRANSACTIONS B: Applications Vol. 37 No. 02, (February 2024) 402-411

International Journal of Engineering

Journal Homepage: www.ije.ir

A Model for Scheduling of Electric Vehicles Charging in a Distribution Network using Multi-agent Model

H. Hamidi*[a], A. Tavassoli[b]

[a] *Department of Industrial, K. N. Toosi University of Technology, Tehran, Iran*
[b] *Department of Information Technology, E-branch, Islamic Azad University, Tehran, Iran*

PAPER INFO

Paper history:
Received 30 August 2023
Received in revised form 28 September 2023
Accepted 30 September 2023

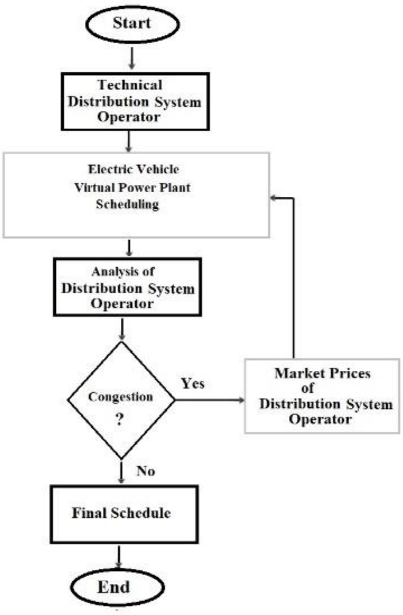
Keywords:
Multi-agent Model
Electric Vehicles
Aggregator
Energy Management System
Bidding Strategy
Electric Market

ABSTRACT

This study has been conducted aiming at improvement of a multi-agent model whose task is planning and energy management of a power distribution system based on electric vehicles and their aggregators. In this work, the wear of automobile batteries is considered as an inhibitor agent for electric vehicle owners which affects other agents. Therefore, the aggregator agent should consider the cost as encouragement for the owners of electric vehicles. The agents used in this paper are: 1) Technical agent distribution system operator 2) Distribution System Operator market agents 3) Electric vehicle aggregator agents. This paper proposes a strategy for the aggregation agent of electric vehicles in a competitive electricity market, taking into account market reservations. This model provides a way to reimburse vehicle owners for battery burnout over the consumption cycle and it helps to increase the desire of electric vehicles to charge and sell electricity to the market and increase the profits of vehicle agents and owners.

doi: *10.5829/ije.2024.37.02b.15*

Graphical Abstarct



---

*Corresponding Author Institutional Email: *h_hamidi@kntu.ac.ir* (H. Hamidi)

Please cite this article as: Hamidi H, Tavassoli A. A Model for Scheduling of Electric Vehicles Charging in a Distribution Network using Multi-agent Model. International Journal of Engineering, Transactions B: Applications. 2024;37(02):402-11.

1. INTRODUCTION

Various methods have been proposed to minimize costs in power distribution networks (1). Today, electricity exchanges in most countries take place through the electricity market and by considering the economic performance of power systems by the system operator, which is required to maintain an acceptable level of security (2). The technology of electric vehicles, or those with battery-powered vehicles, can be divided into two general categories of EVs and Plug-In Hybrid Electric Vehicle (PHEVs). Among these technologies, PHEVs are generally more attended due to the ability to work simultaneously with two internal combustion and electric engines (3, 4).

On the other hand, PHEVs can solve many power system problems as a potential electric energy store and causes the possibility of electric energy storages with high capacity, but sporadic, and random-behavior. Introducing an appropriate aggregator for these types of vehicles and defining its relationships with other power system organizations can help them grow faster. An aggregator is an agent as an interface between the system operator and the consumer. The aggregator helps to facilitate the use of electric vehicles and maximize the profit of operator and vehicles (5). The existence of batteries in these vehicles as their main characteristic, as well as the ability to charge and discharge electricity power through the global power grid lead to change the attitude to these vehicles from an electrical load to a moving electrical energy storage. Managing battery charging and discharging electric vehicle batteries is one of the most important issues in electric vehicle-based smart grids. In fact, the management of vehicle charging and the timing of consumed electricity traffic should be such that the profit of all market participants reaches its optimal level and reduces network costs (6). One of the most basic needs for a healthy competition is the monitoring on parties' performance by the neutral factor. In the electricity market, this task is performed by the Independent System Operator (ISO). Independent system operator is responsible for maintaining the security of the system. This operator acts in market environment independently and does not discriminate between them in encouraging or penalizing participants. It is at the same time responsible for operating the power system and operating from the ultimate market. Generation companies (Gencos) produce and sell electrical energy. They also provides additional ancillary services such as frequency regulation, voltage control and reserve for which the operator needs. A generation company can own one or more power plants with different technologies. Generation companies may also be called Independent Power Producers (IPPs). Transmission companies (Transcos) are the owners of transmission line equipments. These companies use their own equipment in accordance with the independent operating system instructions. Distribution companies (Discos) are the owners and operators of the distribution network. In this paper, the focus is on production and transmission and, distribution network is not used. Virtual power plants are agents of power system without power generation; but they can store electrical power at some times and return it to the grid at other times. In this work, electric vehicles are considered as virtual power plants.

In this paper, a way to optimize the planning of electric vehicle charging is proposed. The proposed model will be a complete model of smart grids that will include all the participants. Participants in the proposed model will include electric vehicles, distributed generation resources and aggregators. Each of the existing agents is trying to maximize its profits in the electricity market. Therefore, various agents must be able to estimate market conditions by predicting load and prices, and present their best offerings for market participation; this is done with the help of a hierarchical multi-agent control system and CPLEX optimization methods with the GAMZ software.

The second section focuses on theoretical foundations and provides the background to the proposed research. The third section, the proposed model is presented. In section 4, the results obtained are analyzed. In section 5, we will present the discussion. in section 6, the conclusion is presented.

2. RESEARCH BACKGROUND

Multi-agent system models were first introduced and used to control the components of smart grids (7), which are most commonly used for large area grids. An example is the use of these models in intelligent energy markets (8-16). The multi-agent system is a new computational model in which several inputs interact with other inputs in an environment (10, 17).

The application of multi-agent systems is divided into two groups: solving the problems of large power systems with the goal of reducing computations (11, 12, 18, 19). Controlling power systems by dividing the system into smaller subsystems and increasing decision-making power (13, 14, 20).

A multi-agent system model is discussed in literature [15-18,21] for managing a distribution system comprising a large number of electric vehicles. The main reason is the use of the power system and, this agent modifies the program with the control and study of the network and informs them.

Hamidi and Moradi (19) presented a multi-agent system is where the charging time of electric vehicles is determined in a way that occurs at the lowest cost of electricity. In this work, the capacity of the transmission lines has been considered as a limiting factor for decision making. It should be kept in mind that the transmission lines should not be loaded more than their capacity.

Hamidi et al. (21) also provided a way to charge vehicles when the network price is the lowest. In this work, electricity prices are also affected by the time of charging vehicles. In other words, the time of minimum price is a function of the time of charging vehicles, and the time of charging vehicles should also occur at the lowest market price. With this in mind, the two variables should be optimized in terms of each other.

Hu et al. (22) provided a multi-agent method according to which, in the environment including electric vehicles, the filling of transmission lines is managed and prevented.

Multi-agent systems are used for a variety of electricity network optimizations. The hierarchical agent-oriented control system is shown in Figure 1.

It is assumed in this work that four agents are considered. These agents are as follows:

1. Technical Distribution System Operator (Technical DSO)
2. Market Distribution System Operator (Market DSO)
3. Electric Vehicle Virtual Power Plant (EV VPP)
4. Dispersed Energy Resources (DER)

In Figure 1, the relationship between these agents is completely identified. The task of technical DSO is to control the technical agents of the network, such as filling lines. The market DSO also optimizes system economic affairs. Also, the task of electric vehicles or EV VPP is the control and management of electric vehicles (23-28).

3. MODELLING

In this paper, electric vehicles owners give the responsibility of charging and discharging their batteries to the aggregators. Aggregators can charge batteries at low power prices and in the high-load hours, additional and unused charge of these vehicles will bring back to the grid. The battery charging by the grid is called Grid-to-Vehicle (G2V), and the return of energy from the vahicle's battery to the grid is called Vehicle-to-Grid (V2G). The owner of any electric vehicle has the ability to or not to participate in the market with the binary variable $\alpha_{t,v} \in \{1,0\}$. If the vehicle is capable of charging and discharging, this variable will show the number 1 and if there is no ability for any reason, such as failure or absence in the parking, then the corresponding variable represents the zero number. This variable specifies the state of the device v at t.

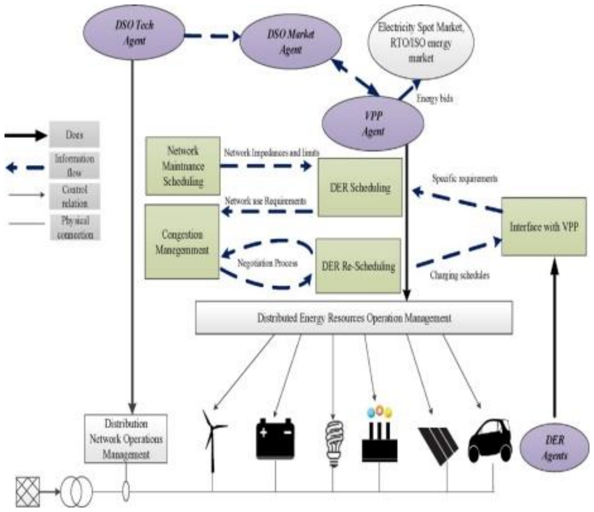


Figure 1. The relationship between various agents in the distribution system (22)

If the EV is ready for charging or discharging, it must behave according to the program that the aggregator specifies and, if necessary, receives energy from the grid, and in other times it will provide energy to the grid. This will increase the owner's profit, electric vehicle aggregator as well as the power system.

The aggregator maximizes its profits by planning electrical vehicles. The aggregator profit is obtained from difference in revenue from services provided to the grid and costs incurred on the grid. One of the charges made in the network is the cost of wearing out electric vehicle batteries. In this work, it is assumed that the aggregator participates in both the regulation up and regulation down market and offers proposals in both. If aggregator offer are accepted by Independent System Operator (ISO), it is necessary to participate in the market and provide the services required by the Independent System Operator. Otherwise, it will suffer damage. Figure 2 shows the decision tree made by the aggregator in the day ahead market. It can be better to see the probabilities in Figure 2. The aggregator offering in the market is that it offers two numbers, one for the price, and one for the market share. To accept the aggregator proposal, a probability is defined which is shown with π^{a}. Also, the probability of deployment in the market is also shown with the variable π^{d}. In the event of winning by the aggregator in the market, two powers are defined; One P^{cap}, which means the aggregator suggested power that specifies the proposed capacity and one P^{accept}, which means its accepted power in day ahead regulation market by ISO.

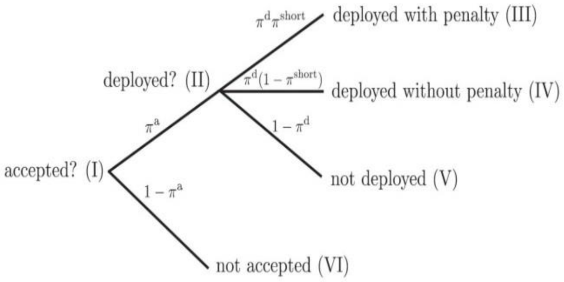


Figure 2. Decision tree for market interactions

In branch (I), the capacity offer is accepted with probability π^a. In the following, and after admission to the market, the potential established by SO may be ranged from zero to p^{accept}, in this case:

$$0 \leq p^{depl} \leq p^{accept} \quad (1)$$

This happens with the probability π^d and is indicated in Figure 2. Also, the aggregator must realize that the power demanded by the SO can be higher than the expected power, and therefore the shortened power or p^{short} must be considered. An ideal case without a penalty factor has been shown in (a) from Figure 3. In this aggregator, the λ^{cap} is established for the accepted capacity p^{accept} and λ^{RT} is paid for p^{depl}.

In branch (V), the aggregator expects that his proposal to be accepted with probability π^a, but not deployed in the RT market. In this case, no gain is achieved, and only the offer placed on the market next day is received at the price of λ^{cap}. In branch (V), the offer is accepted and it is expected that all accepted capacity is not used in RT market. In this case, there is no benefit to participation in the RT market, and not the risk for the aggregator. In fact, in this case, only the offer to participate in the day ahead market and with the price of λ_{cap} is accepted.

In this model, an aggregator which decides to maximize its profit is considered. The objective function of this problem is written as follows.

$$\max\{r^{em} + r^{cap} + r^{depl} - c^{regup} - c^{regdn} - c^{deg}\} \quad (2)$$

In relation 2, the symptoms are as follows:

r^{em} *: Profit from the participation on the day ahead energy market*

r^{cap} *: Profit from the participation on the day ahead energy market*

r^{depl} *:The expected profit in the real time market for up and down regulations are respectively obtained with the probabilities π^a and ϕ^a. Also, the deployment and implementation of a proposal offered in the RT market occurs with probabilities π^a and ϕ^a.*

c^{regup} *: The cost of up regulation services*

c^{regdn} *: The cost of down regulation services*

c^{deg} *: Battery Depreciation Costs*

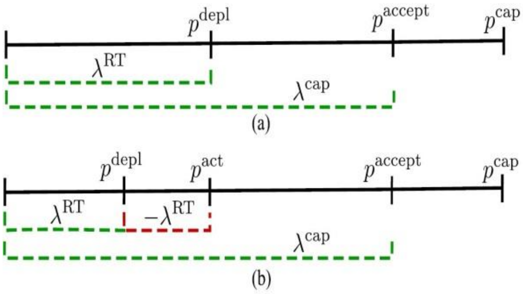


Figure 3. Actual profit and costs for a) a non-penalty b) penalty cases

In the model proposed in this paper, it is assumed that the energy and reserve market are closed by the system operator based on the power pool market. The system operator identifies generators' produced power program by doing Unit Commitement (UC) program. In this work, the energy and reserve market is closed in two steps. In the first stage, the energy market is implemented and the reserve market will be exploited in the second stage.

After closing day ahead market, during the operation of the system, it may happen that there is a need for more or less planned energy. For this reason, there is a need for interaction of RT market. If the required power is greater than the planned power in the DA market, up regulation is required and unsupplied energy is provided. Also, if the load is less than production, this excess power will be sold at up regulation.

In this work, the market price is provided in step form and the price and quantity of power accepted in market are also in step form. Each of the power values and prices offered to them is accepted on the market with different probabilities. Figure 4 shows the total of this process and the final ordered pair of value-price.

It is necessary to define its various variables in order to obtain the objective function,. At first, r^{em} or profit from the DA market is obtained.

The relation 3 of the profit from the day ahead market is as follows:

$$r^{em} = \Delta t \sum_{t \in T} \sum_{v \in V} \lambda_t^{DA} \left(\eta_v^{dsg} . p_{t,v}^{emdsg} - p_{v,t}^{emchg}\right) \quad (3)$$

in relation to 3:

λ_t^{DA}	The electricity price at the hour t of the day ahead market
η_v^{dsg}	Battery discharge efficiency v
$p_{t,v}^{emdsg}$	Battery dischargable power v at hour t

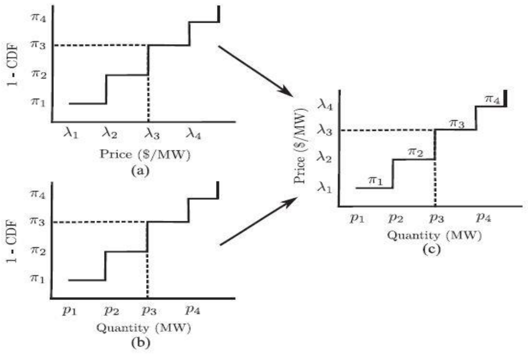


Figure 4. Capacity and price accepted in the market (a) Prices and acceptance probabilities for each price (b) Power values and acceptance probabilities for each value (c) Ordered pair of power- price, which is obtained from the combination of two graphs (a and b)

$p_{v,t}^{emchg}$	Power stored in battery v at hour t
Δt	Study interval for calculating profit

Also, the day ahead regulation market profit is calculated according to relation 4:

$$r^{cap} = \sum_{t\in T} \sum_{b\in B}\left[(\omega_{t,b}^{up}\lambda_{t,b}^{up})\pi^{a}p_{t}^{up} + (w_{t,b}^{dn}\lambda_{t,b}^{dn})\Phi^{a}p_{t}^{dn}\right] \tag{4}$$

In relation 4, the values are as follows:

$\omega_{t}^{up\backslash dn}$	:	A binary variable that indicates whether step b of price-value-probability curve is active or not.
$\lambda_{t,b}^{up\backslash dn}$	:	Regulation price at time t and step b.
π^{a}	:	The acceptance probability (winning) in the regulation up market
Φ^{a}	:	The acceptance probability (winning) in the regulation down market
$p_{t}^{up\backslash dn}$	:	Regulation power suggestion at hour t

The mentioned relationship involves only profits from capacity market commitment. If the system operator needs to use the offered offer in the capacity market, it is necessary to buy at RT market price.

The profit from this work is in the form of a relation 5:

$$r^{depl} = \pi^{a}\pi^{d}\eta^{dsg}\sum_{t\in T} \sum_{v\in V} \sum_{b\in B} (v_{t,b}^{up}\lambda_{t,b}^{RT})(e_{t,v}^{regup} + e_{t,v}^{stopdsg}) \tag{5}$$

In relation 5, $\lambda_{t,b}^{RT}$, the RT market price is obtained from the curve of Figure 4. $e_{t,v}^{regup}$ is also the energy that is expected to be deployed and used in the market for regulation up. Also, $e_{t,v}^{stopdsg}$ is the energy is scheduled for energy discharging in the regulation up service. The r^{depl} benefit discussed in this section is shown in branch (III) of Figure 2. The aggregator must install a part of the capacity offer in order to obtain this benefit, regardless of the risk of consumption more than the proposed power. This will likely result in the aggregator being forced to pay the costs due to lack of capacity. This cost is shown in relation 6.

$$C^{regdn} = \Phi^{a}\Phi^{d}(1-\Phi^{d})\sum_{t\in T} \sum_{b\in B} (v_{t,b}^{up}\lambda_{t,b}^{RT}) (p_{t}^{dn} - \Phi^{a}\Phi^{d}p_{t}^{dn}) \tag{6}$$

$$C^{regup} = \pi^{a}\pi^{d}(1-\pi^{d})\sum_{t\in T} \sum_{b\in B} (v_{t,b}^{up}\lambda_{t,b}^{RT}) (p_{t}^{up} - \pi^{a}\pi^{d}p_{t}^{up}) \tag{7}$$

In relation 7, the difference between the suggested capacity p_{t}^{up} and the power expected to be used indicates the power shortage. This power shortage should be purchased by an aggregator on the RT market. The probability of this power shortage is $\pi^{a}\pi^{d}(1-\pi^{d})$. The term $(1-\pi^{d})$ is called the power shortage probability or π^{short}. This probability has been shown in branch (IV) of Figure 2.

Now, if the power purchased is more than used or deployed, the excess amount should be sold in the RT market. This is called regulation down. Charging and discharging the vehicle batteries will cost the owners of electric vehicles. As a result, they do not want to adjust the charge and discharge of vehicles by the aggregator. Therefore, in order to encourage this, the aggregator should consider a cost as encouragement for them. This amount should be such as to compensate the wear of the batteries due to charge and discharge. Relation 8 is used for wear of the battery of existing electric vehicles:

$$c^{deg} = \sum_{v\in V} C_{v}^{bat}\left|\frac{m_v}{100}\right|\left[\frac{\Delta t\sum_{t\in T}\left(p_{t,v}^{emdsg}+p_{t,v}^{emchg}\right)-\xi_v}{BC_v} + \frac{\sum_{t\in T}\left(\pi^{a}e_{t,v}^{regup}+p_{t,v}^{emchg}\right)}{BC_v}\right] \tag{8}$$

In the relationship 8:

BC_v	Battery capacity of vehicle v
C_v^{bat}	Battery cost of vehicle v
m_v	The linear estimate of battery life of vehicle v based on the number of charge and discharge cycles
ξ_v	Total energy for movement of vehicle v

In relation 8, the aggregator is obliged to compensate the owners of electric vehicles, to transfer energy to the network and wear of the batteries. This amount of compensation is dependent on the discharge power shown with $p_{t,v}^{emdsg}$. For excess energy received from the grid, compensation should also be considered. But this compensation is paid for the excess energy on the use of the vehicle for moving. For example, if energy ξ_v is required for a vehicle to move, wear compensation for battery is considered for the energy that is more than this charge. It should be noted that for regulation services, the energies $e_{t,v}^{regdn}$ and $e_{t,v}^{regup}$, which cause the wear of the batteries, are considered and the number of charge times does not matter and it does not burn out the battery of the electric vehicle. The objective function mentioned at the beginning of the modeling has constraints and limitations that are discussed below.

The relations 9 and 10, respectively, indicate the proposed capacity of regulation up and down.

$$p_{t}^{up}\pi^{d}\Delta t = \sum_{v\in V} (e_{t,v}^{regup} + e_{t,v}^{stopchg}) \quad \forall t \in T \tag{9}$$

$$p_{t}^{dn}\Phi^{d}\Delta t = \sum_{v\in V} (e_{t,v}^{regdn} + e_{t,v}^{stopcdn}) \quad \forall t \in T \tag{10}$$

Now, if the extra capacity of the batteries is taken into account, the aggregator can, by utilizing and scheduling, perform the regulation up and down and do not have to buy the RT market. The constraints of this work are shown in relations 11 to 12:

$$soc_{t,v} = soc_{t-1,v} + \eta_v^{chg} p_{t,v}^{emchg} \Delta t - p_{t,v}^{emdsg} \Delta t - \xi_v \left(\frac{S_{t,v}}{\sum_{t \in T} S(t,v)} \right) \; \forall t \in T, v \in V \qquad (11)$$

$$soc_{t,v} \geq \xi_v \left(\frac{S_{t,v}}{\sum_{t \in T} S(t,v)} \right) \quad \forall t \in T, v \in V \qquad (12)$$

$$0 \leq e_{t,v}^{regdn} + e_{t,v}^{stopdsg} \leq SoC_v^{Max} - soc_{t,v} \quad \forall t \in T, v \in V \qquad (13)$$

$$0 \leq e_{t,v}^{regup} + e_{t,v}^{stopchg} \leq soc_{t,v} - SoC_v^{Min} \quad \forall t \in T, v \in V \qquad (14)$$

In all of the equations and relationships mentioned above, it should be noted that the capacity of the batteries is less than the minimum and does not exceed the maximum. Therefore, the relationship 15 for the batteries is considered:

$$SoC_v^{Min} \leq soc_{t,v} \leq SoC_v^{Max} \quad \forall t \in T, v \in V \qquad (15)$$

A constraint should also be taken into account that to put the battery energy at the end of the day at the constant amount at the beginning of the day. This constraint is considered in relation 16:

$$soc_{t=|T|,v} = SoC_v^{init} \qquad v \in V \qquad (16)$$

4. ANALYSIS

Given the hypotheses of the problem, a unique mathematical model was identified for the problem. Finally, mathematical relations were extracted and formulated according to the proposed model. According to the proposed model, its results are analyzed using the GAMS software and using the CPLEX method. In this work, the system consisting of 100 electric vehicles have been considered which are controlled by an aggregator. The system is equipped with aggregating agents for electric vehicles or virtual power plants (VPP), electric power consumers, electric vehicles and system operators. Electric vehicle information is assumed as follows:

Battery capacity of each vehicle is 24 kWh. The maximum battery charge is 95% and its discharge depth is 15%. The internal energy of the batteries is initially considered as random numbers for each battery. The battery wear rate or gradient, or m_v for each battery, is considered to be -0.015. A similar step curve is used for market offerings. The probabilities π^a and Φ^a equal to 0.9 and the probabilities π^d and Φ^d equal to 0.8. The curve used in this study has four steps, which are represented by the set b. The stepwise prices for all four steps λ^{dn} for 24 hours are shown in Table 1.

TABLE 1. Market price for regulation down

Hours	Step1	Step2	Step3	Step4
1	6	7	8	9
2	6	7	8	9
3	7	8	9	10
4	7	8	9	10
5	8	9	10	11
6	8	9	10	11
7	8	9	10	11
8	8	9	10	11
9	9	10	11	12
10	10	11	12	13
11	11	12	13	14
12	12	13	14	15
13	13	14	15	16
14	14	15	16	17
15	13	14	15	16
16	12	13	14	15
17	11	12	13	14
18	10	11	12	13
19	12	13	14	15
20	13	14	15	16
21	14	15	16	17
22	12	13	14	15
23	10	11	12	13
24	8	9	10	11

Also λ^{up} or the price of each step for regulation up is shown in Table 2.

TABLE 2. Market price for regulation up

Hours	Step1	Step2	Step3	Step4
1	6	7	8	9
2	6	7	8	9
3	7	8	9	10
4	7	8	9	10
5	8	9	10	11
6	8	9	10	11
7	8	9	10	11
8	8	9	10	11
9	9	10	11	12
10	10	11	12	13

Hours	Step1	Step2	Step3	Step4
11	11	12	13	14
12	12	13	14	15
13	13	14	15	16
14	14	15	16	17
15	13	14	15	16
16	12	13	14	15
17	11	12	13	14
18	10	11	12	13
19	12	13	14	15
20	13	14	15	16
21	14	15	16	17
22	12	13	14	15
23	10	11	12	13
24	8	9	10	11

Regulations up and down at different hours of the day are shown in the Figures 5 and 6.

Regulation down energy is shown negatively in Figure 6. Also, in 24 hours a day and for all electrical vehicles in the grid, the sum of the energies $e^{stopchg}$ and total energies $e^{stopdsg}$ are summarized in Table 3.

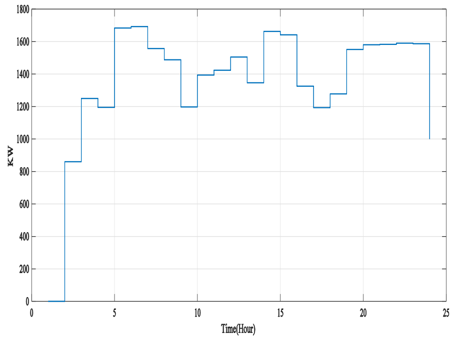


Figure 5. Regulation up energy

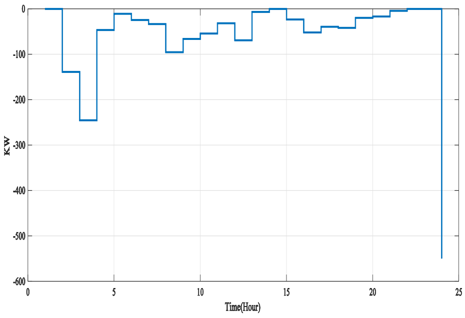


Figure 6. Regulation down energy

TABLE 3. Energies $e^{stopdsg}$ and $e^{stopchg}$

Energy	Value (KWh)
$e^{stopchg}$	32586
$e^{stopdsg}$	1570

The total aggregator profit is obtained from the algebraic sum of revenues and expenses. Revenues and costs and, ultimately, the profits earned by the aggregator are presented in Tables 4 and 5.

$$\text{Costs - Income = Aggregator Benefit} \tag{39}$$

Using the relationship 39, the profit from participating of aggregator in the market is the numerical value of \$ 2.9 × 10^6, in the reference paper, the total cost which aggregator pays for the regulation market is equal to 2.436×10^6. In the research, the total regulation costs are about 1.33×10^5. The reason for this difference is in three cases. The aggregator can achieve huge profits by using only electric vehicle battery capacity and accurately scheduling battery charging and discharging. In fact, the resources that lead to profits for the aggregator are vehicle batteries which acts as virtual power plants (VPP). Now, if the capacity of these batteries changes, the capacity of the VPPs has actually changed and it is expected that this change will strongly affect the energy exchanges and grid financial transactions and aggregator of electric vehicles.

Table 6 shows the effect of battery capacity on aggregator financial transactions. The effect of the battery capacity on the aggregator's profit is shown in Figure 7. As shown in Figure 7, increasing the capacity of the batteries increases the available power of the aggregator and increases its maneuverability. This increased maneuverability will increase its profits.

TABLE 4. Revenue from aggregator participation in markets

Revenues (Income)	Abbreviation	Value($)
Profit from the participation on the day ahead energy market	r^{em}	$5.2*10^4$
Profit from the participation on the day ahead energy market	r^{cap}	$2.2*10^6$
Income from the deployment of power in the RT market	r^{depl}	$8.1*10^5$

TABLE 5. Costs that the aggregator pays

Costs	Abbreviation	Value ($)
The cost of up regulation services	c^{regup}	$8.7*10^4$
The cost of down regulation services	c^{regdn}	$1.3*10^4$
Battery Depreciation Costs	c^{deg}	$3.3*10^4$

TABLE 6. Effect of battery capacity on aggregator profit

Each vehicle's battery capacity (KWh)	Aggregative profit ($)
12 KWh	$1.97*10^6$
15 KWh	$2.43*10^6$
18 KWh	$2.76*10^6$
21 KWh	$2.89*10^6$
24 KWh	$2.96*10^6$

According to Figure 7, it is clear that with increasing battery capacity, profit will be increased but profit growth will decrease.

5. Discussions

In fact, when the price of electrical energy in the grid is low, the aggregator can charge the available vehicle battery with proper planning and when the price of electricity in the grid is high it sells the excessive energy stored in batteries to the grid. In addition to making big profits, the aggregator can pay the owners of electric vehicles to encourage and to compensate the wear batteries. On the other hand, the network's electricity price is high during peak hours and is low during low hours. However, if vehicles get electricity from the grid at low load hours and they bring this energy back to the grid at peak times, help for peak shaving or flattening the load profile. With this, the costs of investment, operation, repair and maintenance of the network also significantly improve. The participation of electric vehicles in the electricity market depends on their charging status. Table 7 and Figure 8 show the status for all vehicles within the next 24 hours.

How to charge a vehicle has an inverse relationship with power consumption and its price. As shown in Figure 8, during the hours when the electricity price is low (early and late hours), the charge status of vehicles is higher than when the price of electricity is high (mid-day hours). The reason for this is that vehicles prefer to deliver power to the grid in the time of expensive electricity and when it's cheap, buy electricity from the electricity grid. In fact, the management of vehicle charging and the timing of electricity consumption should be such that the profits of all market participants reach their optimum level and reduce network costs. This issue has always been an important topic for electric vehicles and attracted experts' attention.

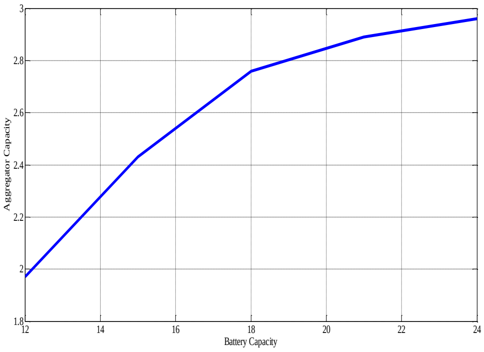


Figure 7. The Effect of Battery Capacity on Aggregator Profit

TABLE 7. Charging status of electric vehicles

hours	Charging vehicles	hours	Charging vehicles
1	1235.667	13	2055.978
2	1840.482	14	1942.5
3	1889.484	15	1861.837
4	2075.37	16	1865.619
5	2043.889	17	1933.543
6	1972.539	18	1998.315
7	1917.485	19	1991.335
8	1867.943	20	1960.806
9	1917.572	21	1922.909
10	1933.761	22	1870
11	1963.472	23	1786.667
12	1944.741	24	1200

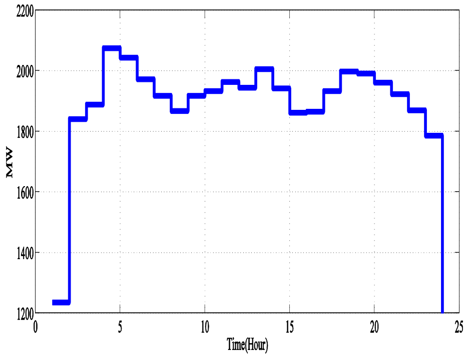


Figure 8. Charging status of electric vehicles

6. CONCLUSION

In this work, the focus was on multi-agent model technology for power management of electric vehicle systems. In the electricity market, there are agents or participants that facilitate the exchanges and operation of the electricity grid. These factors are usually identified as follows:

- The agent of Distribution System Operator (DSO)

- Market agents of the distribution system operator who are responsible for the economic activity of the system.
- Agents of virtual power plants of electric vehicles (in this case electric vehicles are considered as virtual power plants).
- Electric vehicle agents

Considering the relationship between these agents and market participants has led to a comprehensive model governing the electric vehicle-based markets. The proposed model is a multi-agent model and is a topic that is less of a concern for researchers. For this reason, with the advancement and increase of electric vehicles in the urban structure, there are many problems that, despite the many advantages of electric vehicles, makes it even more difficult to use these vehicles. Therefore, it is necessary to provide a comprehensive and optimal plan for the interaction between the agents and the optimum charging of vehicles considering the complete structure of the multi-agent systems. It can also provide relative safety and welfare for the grid and owners of electric vehicles while optimizes the profitof all market participants.

7. REFERENCES

1. Hamidi H, Moradi Abadi A, Amin Mousavi S. A New Method for Open Government using of Information Technology. International Journal of Engineering, Transactions C: Aspects. 2022;35(3):493-501. 10.5829/ije.2022.35.03C.01
2. 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.
3. Pandžić H, Dvorkin Y, Wang Y, Qiu T, Kirschen DS, editors. Effect of time resolution on unit commitment decisions in systems with high wind penetration. 2014 IEEE PES General Meeting| Conference & Exposition; 2014: IEEE. 10.1109/PESGM.2014.6939548
4. Hamidi H, Vafaei A. Evaluation of fault tolerant mobile agents in distributed systems. International Journal of Intelligent Information Technologies (IJIIT). 2009;5(1):43-60. 10.4018/jiit.2009010103
5. 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: Applications. 2023;36(2):321-34. 10.5829/ije.2023.36.02b.12
6. 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
7. Wang Y, Xia Q, Kang C. Secondary forecasting based on deviation analysis for short-term load forecasting. IEEE Transactions on Power Systems. 2010;26(2):500-7.
8. Zhang Y-j, Ren Z. Real-time optimal reactive power dispatch using multi-agent technique. Electric Power Systems Research. 2004;69(2-3):259-65.
9. Disfani VR, Fan L, Piyasinghe L, Miao Z. Multi-agent control of community and utility using Lagrangian relaxation based dual decomposition. Electric Power Systems Research. 2014;110:45-54. 10.1016/j.epsr.2014.01.009
10. Hamidi H, Kamankesh A. An approach to intelligent traffic management system using a multi-agent system. International Journal of Intelligent Transportation Systems Research. 2018;16:112-24. 10.1007/s13177-017-0142-6
11. Karfopoulos EL, Hatziargyriou ND. A multi-agent system for controlled charging of a large population of electric vehicles. IEEE Transactions on Power Systems. 2012;28(2):1196-204. 10.1109/TPWRS.2012.2211624
12. Miranda J, Borges J, Mendes MJ, Valério D. Development of a multi-agent management system for an intelligent charging network of electric vehicles. IFAC Proceedings Volumes. 2011;44(1):12267-72. https://doi.org/10.3182/20110828-6-IT-1002.03492
13. 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.
14. Unda IG, Papadopoulos P, Skarvelis-Kazakos S, Cipcigan LM, Jenkins N, Zabala E. Management of electric vehicle battery charging in distribution networks with multi-agent systems. Electric Power Systems Research. 2014;110:172-9. https://doi.org/10.1016/j.epsr.2014.01.014
15. 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.
16. Hamidi H, Vafaei A, Monadjemi SAH. Analysis and evaluation of a new algorithm based fault tolerance for computing systems. International Journal of Grid and High Performance Computing (IJGHPC). 2012;4(1):37-51. 10.4018/jghpc.2012010103
17. Rahman MS, Oo A. Distributed multi-agent based coordinated power management and control strategy for microgrids with distributed energy resources. Energy conversion and management. 2017;139:20-32. https://doi.org/10.1016/j.enconman.2017.02.021
18. 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: Applications. 2022;35(5):1082-8. 10.5829/ije.2022.35.05b.23
19. Hamidi H, Moradi S. Analysis of consideration of security parameters by vendors on trust and customer satisfaction in e-commerce. Journal of Global Information Management (JGIM). 2017;25(4):32-45. 10.4018/JGIM.2017100103
20. Van der Veen R, Kisjes K, Nikolic I. Exploring policy impacts for servicising in product-based markets: A generic agent-based model. Journal of Cleaner Production. 2017;145:1-13. 10.1016/j.jclepro.2017.01.016
21. 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.
22. Hu J, Morais H, Lind M, Bindner HW. Multi-agent based modeling for electric vehicle integration in a distribution network operation. Electric Power Systems Research. 2016;136:341-51. 10.1016/j.epsr.2016.03.014
23. Hamidi H, Fazeli K. Using Internet of Things and biosensors technology for health applications. IET Wireless Sensor Systems. 2018;8(6):260-7. 10.1049/iet-wss.2017.0129
24. Mengistu D, Davidsson P, Lundberg L, editors. Middleware support for performance improvement of MABS applications in the Grid environment. Multi-Agent-Based Simulation VIII: International Workshop, MABS 2007, Honolulu, HI, USA, May 15, 2007, Revised and Invited Papers 8; 2008: Springer.
25. Hamidi H, Mohammadi K. Modeling fault tolerant and secure mobile agent execution in distributed systems. International

Journal of Intelligent Information Technologies (IJIIT). 2006;2(1):21-36. 10.4018/jiit.2006010102

26. Daraei A, Hamidi H. An efficient predictive model for myocardial infarction using cost-sensitive j48 model. Iranian journal of public health. 2017;46(5):682.
27. Torabi A, Hamidi H, Safaie N. Effect of sensory experience on customer word-of-mouth intention, considering the roles of customer emotions, satisfaction, and loyalty. International Journal of Engineering, Transactions C: Aspects,. 2021;34(3):682-99. 10.5829/ije.2021.34.03c.13
28. Asadi Saeed Abad F, Hamidi H. An architecture for security and protection of big data. International Journal of Engineering, Transactions A: Basics. 2017;30(10):1479-86. 10.5829/ije.2017.30.10a.08

Persian Abstract

چکیده

هدف از این مقاله، بهبود یک مدل چندعاملی است که وظیفه آن برنامه‌ریزی و مدیریت سیستم توزیع انرژی الکتریکی با محوریت خودروهای الکتریکی و تجمیع کنندگان آن‌هاست. در این کار فرسودگی باتری خودروها به‌عنوان یک عامل بازدارنده، برای صاحبان خودروهای الکتریکی در نظر گرفته‌شده است که عامل‌های دیگر را تحت تأثیر قرار می‌دهد. لذا عامل تجمیع کننده، باید هزینه‌ای به‌عنوان مشوق برای صاحبان خودروهای الکتریکی در نظر بگیرد. عوامل استفاده‌شده در این مقاله عبارت‌اند از:1) عامل فنی بهره‌بردار سیستم توزیع 2) بهره‌بردار بازار توزیع 3) تجمیع کنندگان خودروهای الکتریکی . این مقاله، استراتژی پیشنهاددهی برای عامل تجمیع کننده‌ی خودروهای الکتریکی، در بازار برق رقابتی و با در نظر گرفتن رزروهای بازار را انجام می‌دهد. این مدل، راهکاری برای غرامت دادن به صاحبان خودروهای الکتریکی برای فرسودگی باتری‌ها در طول چرخه‌ی مصرف ارائه می‌کند و کمک می‌کند که تمایل بهره‌برداری از خودروهای الکتریکی برای شارژ و فروش برق به بازار بیشتر شده و سود بهره‌بردار و صاحبان خودروها افزایش یابد.

IJE TRANSACTIONS B: Applications Vol. 37 No. 02, (February 2024) 412-424

International Journal of Engineering

Journal Homepage: www.ije.ir

Genomic Ancestry Inference of Admixed Population by Identifying Approximate Boundaries of Ancestry Change

F. Alizadeh[a], H. Jazayeriy*[a], O. Jazayeri[b], F. Vafaee[c,d]

[a] *Faculty of Electrical and Computer Engineering, Babol Noshirvani University of Technology, Babol, Iran*
[b] *Department of Molecular and Cell Biology, Faculty of Science University of Mazandaran, Babolsar, Iran*
[c] *School of Biotechnology and Biomolecular Sciences, University of New South Wales (UNSW), Sydney, Australia*
[d] *UNSW Data Science Hub, University of New South Wales (UNSW), Sydney, Australia*

PAPER INFO

Paper history:
Received 30 August 2023
Received in revised form 25 September 2023
Accepted 07 October 2023

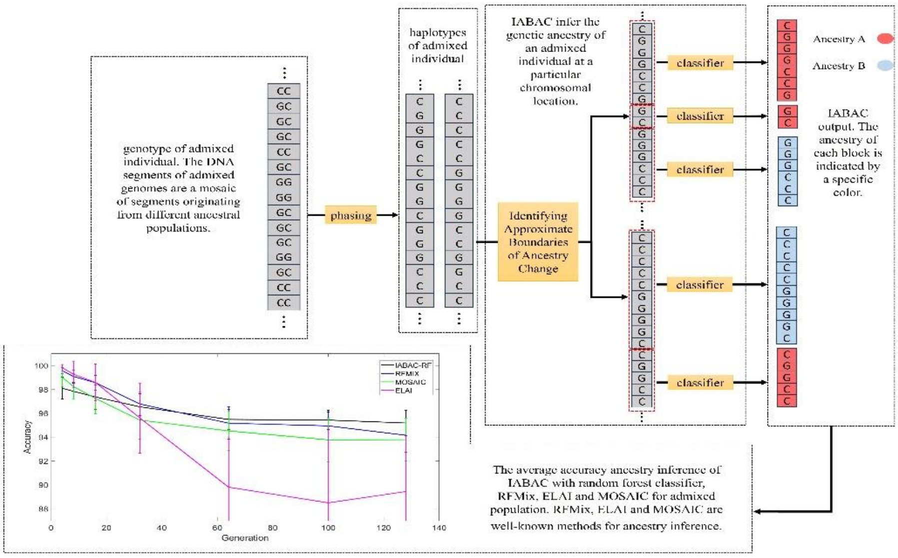
Keywords:
Admixed Haplotype
Admixed Population
Ancestry Inference
Classification
Haplotype Block
Local Ancestry

ABSTRACT

Admixture is a common phenomenon in human populations, resulting from the mating of individuals from two or more previously isolated populations. This can lead to the formation of mosaic DNA segments, with each segment originating from a different ancestral population. Local ancestry inference methods are used to identify the ancestry of each segment, which can provide insights into the history of admixture in a population. Many local ancestry inference (LAI) methods require the determination of various parameters that may be difficult to obtain, which can hamper using LAI methods. In this paper, we present a novel method for identifying approximate boundaries of ancestry change (IABAC) in admixed haplotypes and then determining the ancestry between boundaries. Unlike many LAI methods, our method does not rely on many statistical or biological parameters, therefore more robust to variations in admixture patterns. We evaluate our method on human data, and show that it is more accurate than existing methods for ancestry detection. Our results suggest that IABAC is a promising new method for identifying ancestry boundaries in admixed haplotypes. This method could be used to study the history of admixture in human populations, and to identify genetic variants that are associated with different ancestral populations.

***doi**: 10.5829/ije.2024.37.02b.16*

Graphical Abstract



---

*Corresponding Author Email: *jhamid@nit.ac.ir* (H. Jazayeriy)

Please cite this article as: Alizadeh F, Jazayeriy H, Jazayeri O, Vafaee F. Genomic Ancestry Inference of Admixed Population by Identifying Approximate Boundaries of Ancestry Change. International Journal of Engineering, Transactions B: Applications. 2024;37(02):412-24.

1. INTRODUCTION

Genetic diversity in the DNA sequences of humans is the result of inheritance processes, including mutation and recombination (1). When two or more previously isolated populations mate, the resulting offspring are admixed, meaning that their genomes contain DNA segments from both populations (2, 3). This admixture creates new genetic recombination breakpoints, which can lead to the formation of diverse genomes with mixed DNA segments. The DNA segments of admixed genomes are a mosaic of segments originating from different ancestral populations (4-10).

As travel around the world becomes easier, admixed populations and their complexity are increasing. This is because more and more people are having contact with people from other populations, which can lead to interbreeding and the formation of new admixed populations (4). Figure 1 shows that how the genomes of new populations are created from their ancestors. The chromosomes of more recent generations are a mosaic of ancestral chromosomes.

The relationship between genetic variation and disease risk can vary between ancestral populations. This is because different ancestral populations have different allele frequencies at specific genetic variants (11). Ancestry inference is the process of determining the ancestral populations that contributed to an individual's genome. This is important for a variety of applications, including pharmacogenomics and the study of human demography (2, 6, 12-14).

The availability of genotype and haplotype data has made it possible to statistically infer the admixture history of human populations (4). Several computational methods and tools have been developed for this purpose. One approach, known as local ancestry inference (LAI), identifies the ancestry of each segment of an individual's genome. LAI methods infer the genetic ancestry of an individual at a particular chromosomal location. This information can be used to study the history of human migrations and to identify genetic variants that are associated with specific ancestries. LAI has been used in a variety of models and tools, including SupportMix (15), RFMix (16), and LAMP (17).

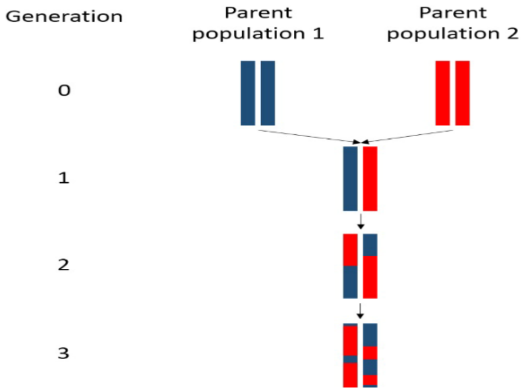


Figure 1. The process of combining the genomes of ancestral populations after generations

Local ancestry inference (LAI) methods subdivide chromosomes into smaller segments, or blocks, to infer the ancestry of each block. The choice of block size is an important factor in the accuracy of LAI. If the block size is too large, it may contain segments from multiple ancestries, which can lead to inaccurate ancestry inference. On the other hand, if the block size is too small, it may not contain enough information to accurately identify the ancestry. The ideal block size should be large enough to contain enough information to identify the ancestry, but small enough to ensure that each block contains only one ancestry (18). This can be a challenging task, as the ancestry of each block can vary depending on the individual's genetic makeup. A number of studies have investigated the optimal block size for LAI. However, the optimal block size may vary depending on the dataset and the LAI method used (17, 19).

Hidden Markov models (HMMs) are a popular approach for local ancestry inference (LAI). HMMs can model the correlation between the ancestries of blocks, which is due to linkage disequilibrium (LD) (15). LD is a phenomenon where genetic variants that are close together on a chromosome are more likely to be inherited together. Some HMM-based LAI methods are SupportMix, PCAdmix (20), MOSAIC (21) and ELAI (22). SupportMix first divides the genome into blocks with a fixed length. Then, it uses a support vector machine (SVM) to determine the ancestry of each block. PCAdmix also divides the genome into blocks with a fixed length. However, it uses a principal components algorithm to determine the ancestry of each block. RFMix determines local ancestry by using a conditional random field (CRF). A CRF is a statistical model that can model the dependencies between multiple variables. RFMix divides the genome into blocks with a fixed length. Then, it uses a conditional random field, parameterized by random forest trained on reference panels, to infer local ancestry within each block. MOSAIC uses nested HMMs to model the correlation between the ancestries of blocks. This allows MOSAIC to infer the ancestry of each segment more accurately than methods that do not use nested HMMs. ELAI employs a two-layer hidden Markov model to obtain local ancestry of each admixed individual.

XGMIX, LAI-NET, LAMP, WINPOP and EILA are also local ancestry inference (LAI) methods that use different approaches to infer the ancestry of each segment of an individual's genome. XGMIX (23) divides the genome into blocks with a fixed length and then infers local ancestry within each window by using gradient boosting trees. The initial estimates are then smoothed using a sliding gradient boosting tree. LAI-NET (24) also divides the genome into blocks with a fixed length, but it uses a neural network model to infer local ancestry within each window. The initial estimates are then smoothed afterwards. Similarly to XGMIX and LAI-NET methods, SALAI-Net (25) follows a two-stage approach: a reference matching layer and then a smoother layer. The reference matching layer infers initial estimates of ancestry for each block, and then the smoother layer improves the initial estimates of ancestry by using neighbouring block information and smoothing ancestry. Wang et al. (26) first divides the genome into blocks with a fixed length and then clusters ancestry within each window by use of localized haplotype clustering (27). The initial estimates are then smoothed using the HMM. LAMP and WINPOP (19) first find the optimal window length on the basis of recombination events and then use a clustering algorithm to find ancestral populations. In these methods, the number of recombination events over time is considered as a Poisson distribution. EILA (28) uses fused quantile regression to identify breakpoints of the ancestral haplotypes. Then, to infer the ancestry of each segment between breakpoints, it utilizes the k-means classifier. machine learning is being increasingly used to analyze genetics data and detect diseases (29-35). As mentioned, some ancestry inference methods also use machine learning to classify haplotype and genotype data.

It is important to note that the accuracy of LAI can be improved by selecting blocks that do not contain any breakpoints. This is because breakpoints can lead to inaccurate ancestry inference (17). LAI methods require the specification of statistical or biological parameters, such as the recombination rate, genetic maps, and average number of generations since admixture. These parameters can affect the accuracy of the LAI results. Among LAI methods, RFMix has been shown to have high accuracy in inferring admixed individuals with two and three ancestral populations.

In this study, we introduce a method called IABAC (Inferring Ancestry using Boundaries of Ancestry Change) that first infers the approximate boundaries of the ancestry change based on the distance between ancestral populations. Then, IABAC identifies the ancestry between boundaries, which are called haplotype blocks. LAI methods require various parameters to be determined, which can make LAI practical uses difficult. It is usually difficult to access these parameters. For example, RFMix requires a genetic map, a window size and the average number of generations since admixture, MOSAIC requires a recombination rates files and SNP files, ELAI requires a SNP position file and the number of upper and lower layers of clusters. Unlike many LAI methods, IABAC does not required to many statistical or biological parameters. the only input parameter of IABAC is the length of the IB-block, which the optimal value of it for ancestry inference is investigated in the next sections.

SupportMix, PCAdmix, RFMix, XGMIX, SALAI-Net and LAI-NET unlike IABAC divide the genome into blocks with a fixed length and determine the ancestry of each block. LAMP and WINPOP first find the optimal block length on the basis of recombination events and then find the ancestral population between two recombination events. LAMP and WINPOP need parameters such as recombination rate to find recombination events. EILA, like IABAC, identifies boundaries of ancestry change. The main difference between EILA and IABAC is that EILA uses fused quantile regression to identify boundaries of ancestry change and IABAC uses distance between ancestral populations to identify boundaries of ancestry change. EILA is used for genotype data and IABAC is used for haplotype data.

We used four classification methods to identify the ancestry of each haplotype block: decision tree (DT), support vector machine (SVM), random forest (RF), and logistic regression (LO). We named IABAC with four different classifiers (IABACs) as IABAC-SVM, IABAC-DT, IABAC-RF, and IABAC-LO, respectively.

We compared the accuracy of the ancestry detection of admixed individuals by IABAC with the fixed window method. In the fixed window method, haplotypes were divided into blocks with a fixed length.

Finally, we compared the performance of IABACs with three well-known benchmark methods: RFMix, ELAI, and MOSAIC.

2. MATERIALS AND METHODS

In this section the research method for identifying the ancestry of admixed individuals is presented.

2. 1. Identifying Boundaries of Ancestry Change

To identify the boundaries of ancestry changes, we consider the allele frequencies of single-nucleotide polymorphisms (SNPs). This method calculates the distance between the alleles of an admixed individual by taking the mean of the alleles of the ancestral populations in a number of predefined SNPs. This distance is denoted by D. The mean of the alleles of each ancestral population in an SNP is obtained from Equation 1.

$$\mu = \frac{\sum_{i=1}^{N} h_i}{N} \tag{1}$$

In this equation, μ represents the mean of the alleles of the ancestral population per SNP. It is a value between 0 and 1. h represents the haplotype allele of each individual in each SNP. It is either 0 or 1. N represents the number of individuals in the population.

If the mean of the alleles of the ancestral population A is denoted by μ_a and the allele of the admixed individual is denoted by h_{ad}, the distance between the admixed individual allele and the mean of the alleles of the ancestral population A in a SNP can be obtained from Equation 2.

$$d_a = |(h_{ad} - \mu_a)| \tag{2}$$

In this equation, d_a represents the distance between the admixed individual allele and the mean of the alleles of the ancestral population A.

When examining the existence or non-existence of the boundaries of ancestry change between two desired SNPs, the distance of alleles in one SNP alone does not have enough information. Therefore, several SNPs need to be considered together. The information of the neighbors of the two SNPs can be used to examine the points of ancestry change between the two SNPs. For example, if we want to examine the ancestry change between two adjacent SNPs i and j, where $j > i$, the number of L_w SNPs from the left neighbor i is considered as one block, and the number of L_w SNPs from the right neighbor j is considered as another block.

In this paper, we call these blocks IB-blocks. The selection of IB-blocks is shown in Figure 2. L_w is the number of neighbors of each SNP that can be selected for different sizes. Figure 2a presents IB-blocks with SNPs i and j (SNPs in these IB-blocks are shown in the red square). In Figure 2b, position i to $i-4$ is considered as one IB-blocks and position j to $j+4$ is considered as another IB-blocks. The value of L_w in Figure 2 is 4. The top row indicates the alleles, and the bottom row indicates the location of the SNPs.

After determining the IB-blocks, the distance between the alleles of the admixed individual in each IB-block and the mean of the alleles of the ancestral population equivalent to that IB-block in each ancestral population is calculated. For the left neighbors (left IB-block) of location i, the distance is calculated using Equation 3. For the right neighbors (right IB-block) of location j, the distance is calculated using Equation 4.

$$D_{aL} = \sum_{k=-L_w}^{0} d_{a(i+k)} \tag{3}$$

$$D_{aR} = \sum_{k=0}^{L_w} d_{a(j+k)} \tag{4}$$

The length of IB-block is shown as L_w and D_{aL} represents the distance between the left IB-block of the admixed individual and the ancestral population A, and D_{aR} represents the distance between the right IB-block of the admixed individual and the ancestral population A. The following algorithm is used to check the boundaries of ancestry changes between location i and j of an admixed individual with ancestry population A and B.

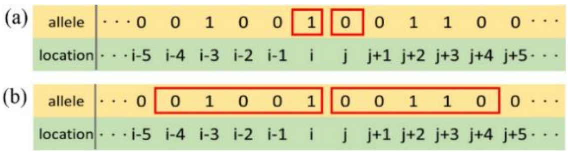


Figure 2. Investigating the existence or non-existence of the boundaries of ancestry change between locations i and j

Step 1: Determine the IB-block and calculate distance D between the admixed individual and the ancestral populations.

Step 2: Select the smallest distance between the ancestral populations and admixed individual for each IB-block. In Equation (5), D_{aR} represents the distance between the admixed individual and mean of the alleles of the ancestral population A in the right IB-block of location j, D_{bR} is the distance between the admixed individual and mean of the alleles of the ancestral population B in the right IB-block of the location j, D_R is the minimum distance between the admixed individual and mean of the alleles of ancestral populations A and B in the right IB-block of location j.

$$\begin{aligned} D_R &= min\,(D_{aR}, D_{bR}) \\ D_L &= min\,(D_{aL}, D_{bL}) \end{aligned} \tag{5}$$

where D_{aL} represents distance between admixed individual and mean of the alleles of ancestral population A in the left IB-block of location i, D_{bL} is distance between admixed individual and mean of the alleles of ancestral population B in the left IB-block of location i and D_L is the minimum distance between the admixed individual and mean of the alleles of the ancestral populations A and B in the left IB-block of location i.

Step 3: Select the smallest distance between the ancestral populations and admixed individual for the total IB-blocks, the blue IB-block specified in Figure 3c, the IB-blocks introduced for both location i and j are considered as one IB-block.

$$D_T = min\,((D_{aL} + D_{bL}), (D_{aR} + D_{bR})) \tag{6}$$

In this equation, D_T represents the minimum distance between the admixed individual and mean of the alleles of the ancestral populations A and B in sum of two IB-blocks of the right and left of location i and j.

Step 4: Compare sum of the minimum distance between the left and right IB-blocks specified in step 2 and the IB-block specified in step 3. If the minimum distance between the total IB-blocks and the value of D_T is not the same, the location between location i and j is selected as the boundary.

$$boundary = \begin{cases} false & D_T = D_R + D_L \\ true & D_T \neq D_R + D_L \end{cases} \tag{7}$$

The steps of determination of the existence or non-existence of the boundaries of ancestry change between locations i and j are shown in Figure 3. In Figure 3a, IB-blocks are determined for locations i and j. The alleles of the admixed individual are displayed with allele AD, the mean of the alleles of the ancestral population A are presented with allele mean A, and the mean of the alleles of the ancestral population B are displayed with allele mean B. In Figure 3b, the distance between the admixed individual and the ancestral population A for each IB-block is shown by the red rectangle, and the distance between the admixed individual and the ancestral population B for each IB-block is shown by the blue rectangle. The distance value for this hypothetical example is shown on the rectangle of each IB-block. An ancestral population with a smaller distance (rectangle) is selected for each IB-block.

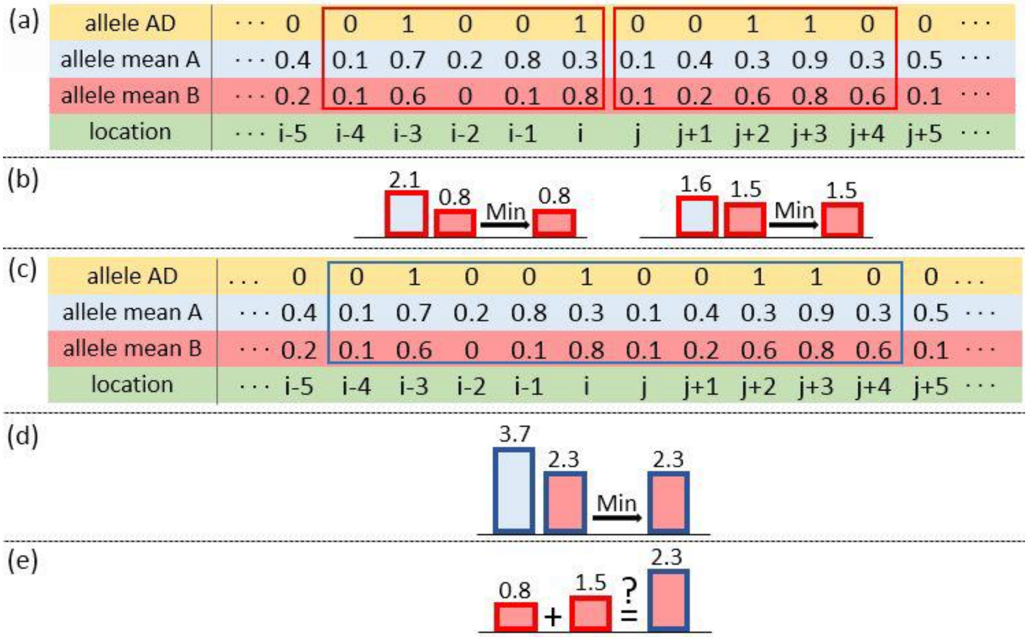
Determination of total IB-blocks are shown in Figure 3c. This IB-block is the sum of the left neighbors for i and the right neighbors for j. In Figure 3d The distance between the admixed individual and the ancestral population A for the total IB-block is shown by the red rectangle, and the distance between the admixed individual and the ancestral population B for the total IB-block is shown by the blue rectangle. An ancestral population with a smaller distance is selected. Figure 3e presents Comparison of the sum of the minimum distance between the left and right IB-blocks specified in Figure 3b and the total IB-block specified in Figure 3d. In this example, there is no boundary of ancestry change between locations i and j, because the sum of the minimum distance between the left and right IB-blocks is equal with the total IB-block.

In the same way, the existence or non-existence of the boundaries of ancestry change between all SNPs are investigated. These IB-blocks are placed as sliding windows between all the SNPs and their distance is calculated. For example, to examine the boundary of ancestry changes between locations j and $j+1$, as shown in Figure 4b, location j to $i-3$ is considered as one IB-block (left IB-block) and position $j+1$ to $j+5$ are considered as another IB-block (right IB-block). The value of L_w in this Figure is 4.

The IB-blocks shown in Figure 4a are the IB-blocks defined to determine the boundary of ancestry changes between location i and j, the IB-blocks indicated in Figure 4b are the IB-blocks defined to determine the boundary of ancestry changes between location j and $j+1$. The left neighbors (left IB-block) of location j and the right neighbors (right IB-block) of location $j+1$ in an IB-block with L_w length for the admixed individual and the ancestry A are obtained from Equations 8 and 9, respectively.

$$D_{aL} = \sum_{k=-L_w}^{0} \mu_{a(j+k)} \tag{8}$$

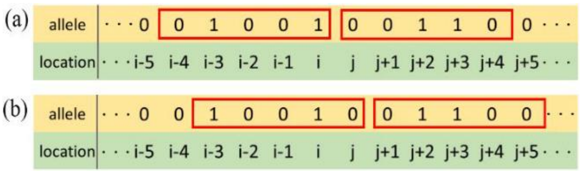
$$D_{aR} = \sum_{k=0}^{L_w} \mu_{a(j+1+k)} \tag{9}$$



(a), (c)

allele AD	···	0	0	1	0	0	1	0	0	1	1	0	0 ···
allele mean A	···	0.4	0.1	0.7	0.2	0.8	0.3	0.1	0.4	0.3	0.9	0.3	0.5 ···
allele mean B	···	0.2	0.1	0.6	0	0.1	0.8	0.1	0.2	0.6	0.8	0.6	0.1 ···
location	···	i-5	i-4	i-3	i-2	i-1	i	j	j+1	j+2	j+3	j+4	j+5 ···

Figure 3. Determination of the existence or non-existence of the boundaries of ancestry change between locations i and j



(a), (b)

allele	···	0	0	1	0	0	1	0	0	1	1	0	0 ···
location	···	i-5	i-4	i-3	i-2	i-1	i	j	j+1	j+2	j+3	j+4	j+5 ···

Figure 4. Investigating the existence or non-existence of the boundaries of ancestry change between locations j and $j+1$

In these equations, D_{aL} represents the distance between the left IB-block of the admixed individual and the ancestral population A, and D_{aR} represents the distance between the right IB-block of the admixed individual and the ancestral population A. Similarly, IABAC can also determine the existence or non-existence of the boundaries of ancestry change between SNPs with more than two ancestral populations.

2. 2. Classification

Once the haplotype IB-blocks with appropriate length have been determined using the IABAC method, the ancestry of each of these haplotype IB-blocks must be classified. We used four well-known classification methods to do this: support vector machine (SVM), decision tree (DL), random forest (FR), and logistic regression (LR). We named the four methods as IABAC-SVM, IABAC-DL, IABAC-RF, and IABAC-LR, respectively.

2. 3. Data and Simulation

We used the genotypes of chromosome 10 from the HapMap project (36), which is a database of genetic variation in humans. The genotypes were phased with SHAPEIT (37), a software that estimates haplotypes from genotype data. Admixed individuals were simulated from these haplotypes using a simple hybrid isolation (HI) model. In this model, all individuals in the first generation can mate with each other, but after that, only admixed individuals from the previous generation can mate (38).

We used data from eight populations in the HapMap project, the names of the populations and their IDs are shown in Table 1. We selected 160 unrelated samples from each population, which resulted in 160 haplotypes for each sample (80 genotypes became 160 haplotypes after phasing). We created admixed individuals by randomly mating samples from the ancestral populations. The probability of recombination in chromosome 10 for each generation was set to 1.8, based on the HapMap data (39). We simulated admixed individuals from their ancestors for 2, 4, 8, 16, 32, 64, 100, and 128 generations.

TABLE 1. populations and their IDs.

Population	ID
Northern and Western European Ancestry	CEU
Toscani in Italia	TSI
Gujarati Indians in Houston, Texas	GIH
Yoruba in Ibadan, Nigeria	YRI
Luhya in Webuye, Kenya	LWK
Maasai in Kinyawa, Kenya	MKK
Chinese in Metropolitan Denver, Colorado	CHD
Han Chinese in Beijing, China	CHB

TABLE 2. admixed populations and their ancestral populations.

Number of ancestral populations	Ancestral Population	Admixed population
Two populations	CEU, CHB	CEU-CHB
	CEU, TSI	CEU-TSI
	CEU, YRI	CEU-YRI
	CHD, TSI	CHD-TSI
	LWK, YRI	LWK-YRI
	LWK, MKK	LWK-MKK
	CHD, GIH	CHD-GIH
	GIH, TSI	GIH-TSI
	LWK, TSI	LWK-TSI
Three populations	CHD, TSI, LWK	CHD-TSI-LWK
	CHD, GIH, TSI	CHD-GIH-TSI
	MKK, GIH, LWK	MKK-GIH-LWK

From the 160 single-population individuals, 140 were selected for training and 20 were used to generate admixed individuals for testing. We simulated 20 admixed individuals by test samples from each pair of ancestral populations. We also simulated 30 admixed individuals by test samples from each triplet of ancestral populations, the names of the admixed populations and their ancestral populations are shown in Table 2.

The simulations were performed with Python and MATLAB software, and each haplotype contained 73,832 SNPs.

3. RESULTS

This section presents the results of the proposed IABAC method. Two important parameters that affect the quality of the IABAC method are the length of the IB-blocks and the choice of classification algorithm. To investigate the effect of IB-block length on the results, we performed ancestry inference for admixed individuals from the CEU-TSI, CEU-YRI, and LWK-MKK populations with different IB-block lengths.

The average accuracy of IABACs for these admixed individuals with 10 samples for each population is shown in Figure 5. Accuracy is measured by the percent of SNPs whose ancestries have been correctly identified. Admixed individuals with 32 generations from the admixed time were considered.

As shown in Figure 5, L_w with value of 100 – 300 SNPs is a good IB-block for all classifiers, and as the number of L_w gets higher or lower, the accuracy of the method decreases. The purpose of IABAC is to provide

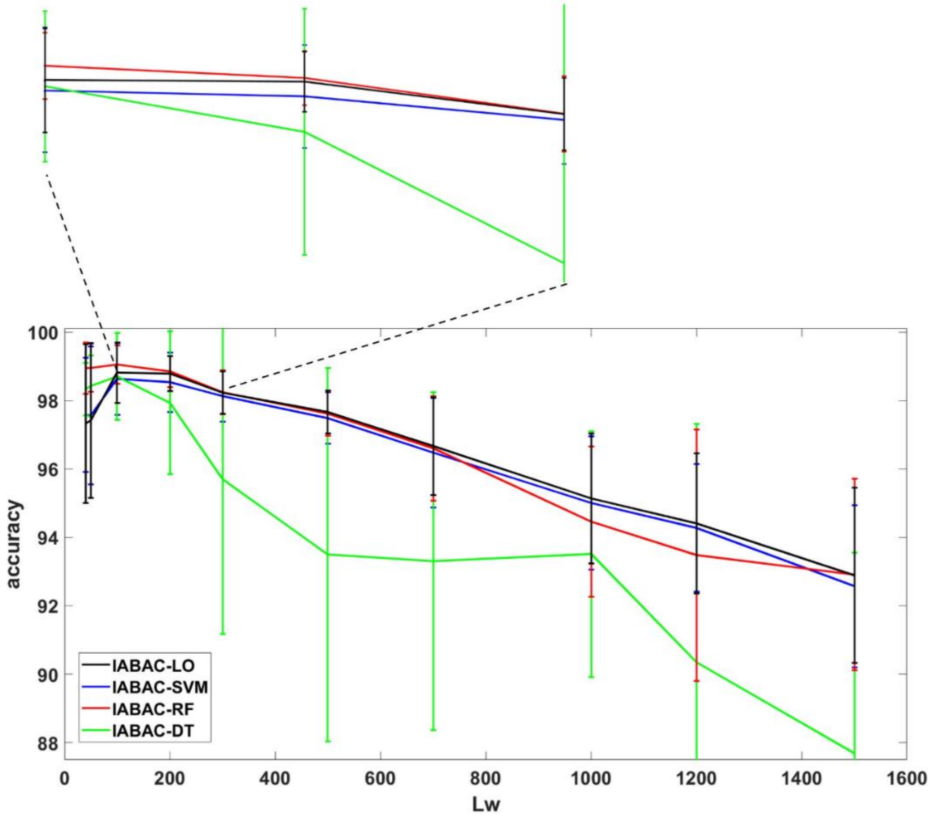


Figure 5. Average accuracy of ancestry inference of admixed individuals with CEU-TSI, CEU-YRI, and LWK-MKK ancestral population in different L_w

an appropriate way to divide chromosomes into smaller haplotype IB-blocks.

To evaluate IABAC and its effect on the accuracy of ancestry inference, the results of IABAC and fixed window with length of 500 SNPs for 20 sample from each admixed individual of CEU-CHB, CHD-TSI and LWK-YRI were compared with each other. The reason for choosing fixed window with length of 500 SNPs is that other ancestry inference methods such as XGMIX (23) and LAI-NET (24) use fixed window with length of 500 SNPs. L_w with a value of 150 SNPs were selected.

Admixed individuals with 2, 4, 8, 16, 32, 64, 100 and 128 generations from the admixed time were considered. The average accuracies of the ancestry detection of admixed individuals using IABAC-SVM and fixed window with SVM classifier (FW-SVM), IABAC-DT and fixed window with DT classifier (FW-DT), IABAC-RF and fixed window with RF classifier (FW-RF), and IABAC-LO and fixed window with LO classifier (FW-LO) are shown in Figure 6.

The results revealed that with increasing admixture times, IABAC-SVM is more accurate than FW-SVM, IABAC-DT is more accurate than FW-DT, IABAC-RF accuracy is better than FW-RF and IABAC-LO accuracy is more valid than FW-LO for ancestry inference. Overall, as the time admixture became longer, the performance of the AICRF algorithm is better than the fixed window in all four classifiers of SVM, DT, RF and LO.

To evaluate IABAC relative to other ancestry inference methods, RFMix, ELAI and MOSAIC are compared with IABACs. Figure 7 presents the average accuracy ancestry inference of IABACs, RFMix, ELAI and MOSAIC for 20 sample from each admixed population of CHD-GIH, GIH-TSI and LWK-TSI (admixed individuals with two ancestral populations). Admixed individuals with 4, 8, 16, 32, 64, 100 and 128 generations from the admixed time are considered. L_w with a value of 150 SNPs is selected.

The results indicated that in low generations ($G < 32$), ELAI performs better than other methods, and with increasing admixture times ($G > 32$), IABAC-RF accuracy is more precise than other methods.

The average accuracy ancestry inference of IABACs, RFMix, ELAI and MOSAIC for 30 sample from each admixed populations of CHD-TSI-LWK, CHD-GIH-TSI and MKK-GIH-LWK (admixed individuals with three ancestral populations) is shown in Figure 8. As shown in Figure 8, in low generations ($G < 16$), ELAI is more accurate than other methods, and with increasing

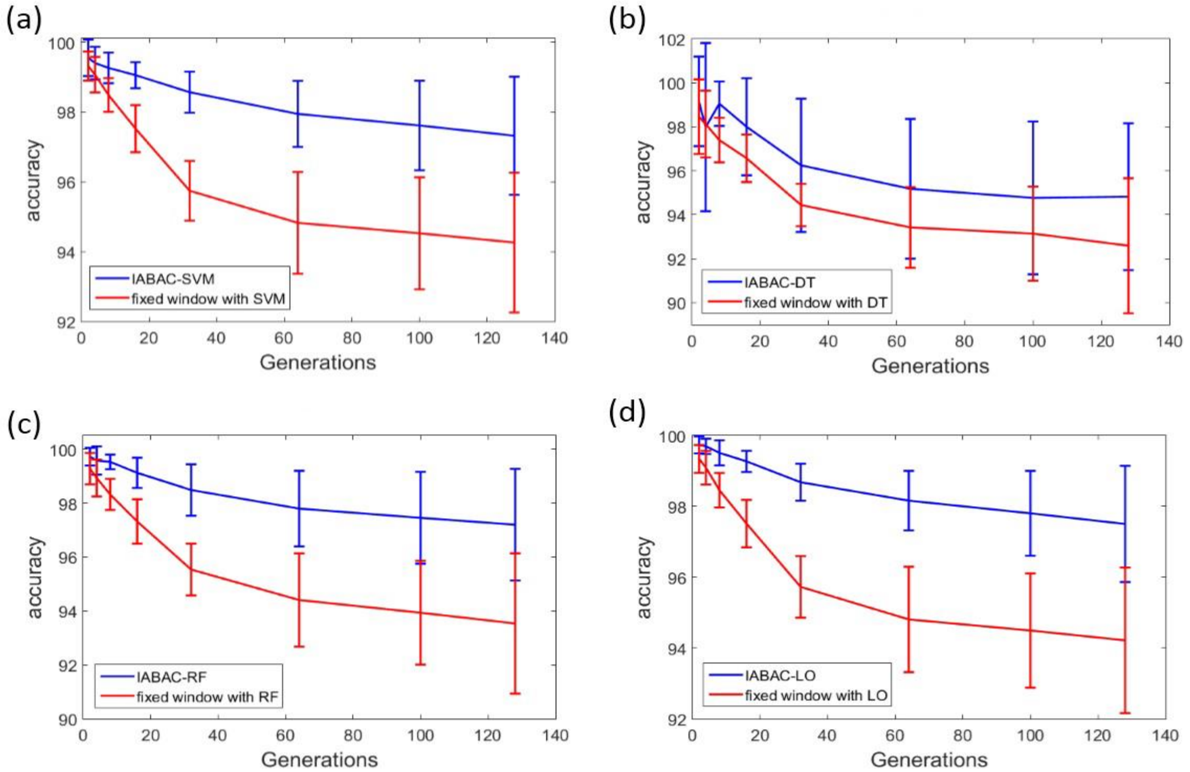


Figure 6. Average accuracy of ancestry inference of admixed individuals using IABAC and fixed window. (a) The average accuracy of ancestry inference using IABAC-SVM and FW-SVM. (b) The average accuracy of ancestry inference using IABAC-DT and FW-DT. (c) The average accuracy of ancestry inference using IABAC-RF and FW-RF. (d) the average accuracy of ancestry inference using IABAC-LO and FW-LO

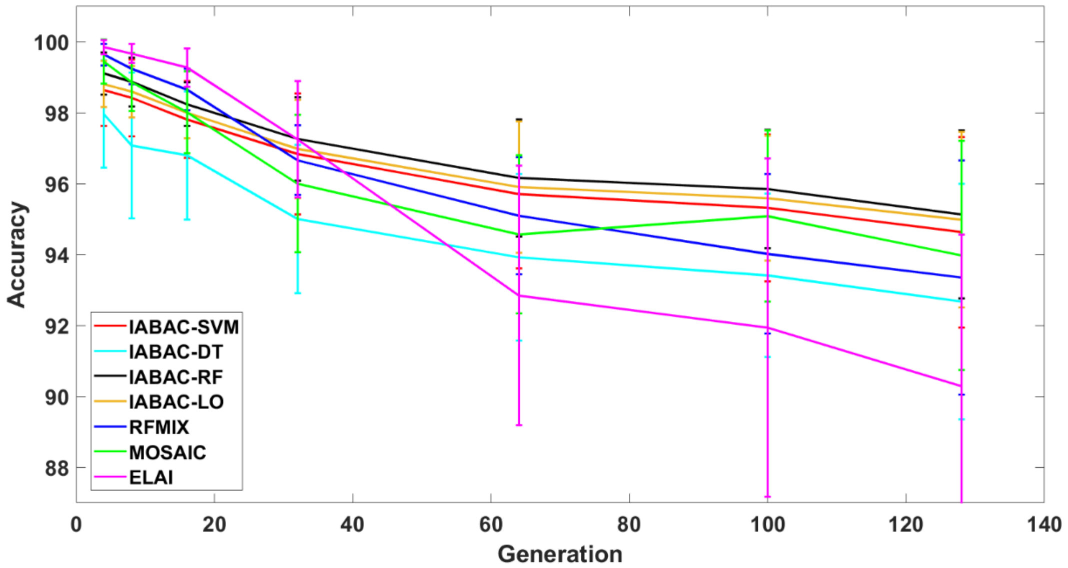


Figure 7. the average accuracy ancestry inference of IABACs, RFMix, ELAI and MOSAIC for 20 sample from each admixed population of CHD-GIH, GIH-TSI and LWK-TSI

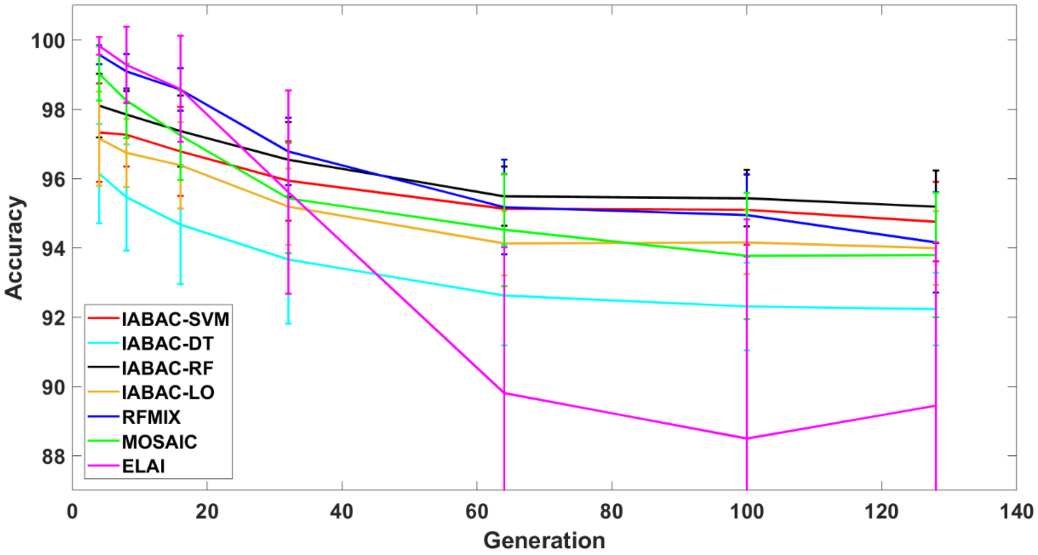


Figure 8. The average accuracy ancestry inference of IABACs, RFMix, ELAI and MOSAIC for 30 sample from each admixed population of CHD-TSI-LWK, CHD-GIH-TSI and MKK-GIH-LWK

admixture times (G > 64), IABAC-RF accuracy is better than other methods.

The results show that with increasing admixture time (G > 64), IABAC-RF can more accurately identify the ancestors of admixed individuals with two and three ancestral population.

An example of an admixed individual of CEU-CHB with its original ancestral population and estimates of IABAC-RF is shown in Figure 9. Admixed individuals with 32 generations from the admixed time are considered. L_w with a value of 150 SNPs are selected. Figure 9a is shown true ancestry of admixed individual, the red blocks represent the ancestral population of CEU and the blue blocks represent the ancestral population of CHB. Figure 9b is presented ancestry estimates of admixed individual, and Figure 9c is indicated difference between true ancestry and estimated ancestry that shown by the red block. The Y-axis represents the probability that one allele is derived from a specific ancestry and the possibility of error in any SNP; the X-axis indicates the physical locations of SNPs. The results of Figure 9 show that most errors occur at ancestry change boundaries. Additionally, some narrow ancestral mosaics have not been correctly detected.

One of the important and influential factors in the accuracy of ancestry inference using IABAC is the chromosome length of the ancestral population constituting the admixed individual chromosome (ancestral mosaics). To examine the effect of the length of the ancestral mosaics, we simulated new admixture populations of CEU-CHB and LWK-YRI. In these admixture people, the ancestral mosaics are equal in length and are repeated alternately.

The results of average accuracy of ancestry inference using IABAC-RF for 20 samples from each admixed population of CEU-CHB and LWK-YRI with different lengths of mosaics are shown in Figure 10. L_w with a value of 150 SNPs were selected.

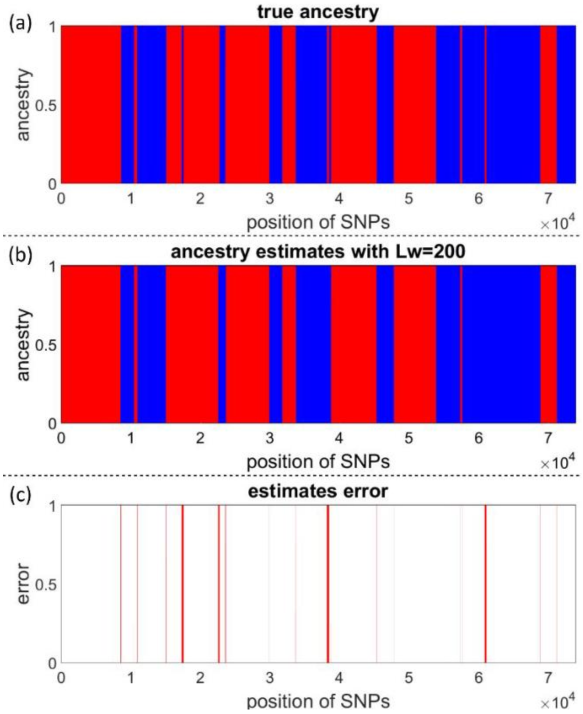


Figure 9. An example of an admixed individual of CEU-CHB with its original ancestral population and estimates of IABAC-RF

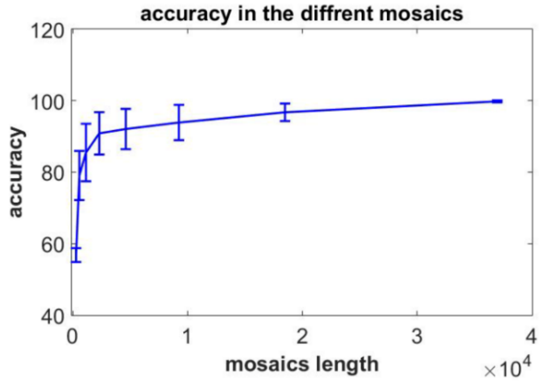


Figure 10. Average accuracy of ancestry inference using IABAC-RF for 20 samples from each admixed population of CEU-CHB and LWK-YRI with different lengths of mosaics

The results show that as the length of ancestral mosaics decreases, the accuracy of ancestry inference decreases. As the admixing time increases, the length of the ancestral mosaics constituting the admixed individual chromosome decreases, so with increasing admixing time, the accuracy of the IABAC-RF decreases.

4. DISCUSSION

Inferring the ancestry of admixed individuals is used in fields of historical demographic, anthropological, and pharmacogenomics. We present a method that uses the distance between the haplotype of ancestral populations to identify boundaries of ancestry change and then determining the ancestry between boundaries. The purpose of IABAC is to provide an appropriate way to divide chromosomes into smaller haplotype blocks. The major strength of IABAC is that it identifies approximate boundaries of ancestry change and ancestry inference without the need to determine many statistical or biological parameters. Many LAI methods require the determination of various parameters that may be difficult to obtain, which can hamper using LAI methods. For example, RFMix requires a genetic map, a window size and the average number of generations since admixture, while the only input parameter of IABAC is the length of the IB-block (L_w).

The results of this study show that the length of IB-blocks in Equations 3 and 4 are influential in the IABAC method. As the length of IB-blocks increases, the accuracy of the IABAC also increases. However, with excessive growth of the length of IB-blocks, the accuracy of the IABAC decreases. The reason for the decrease of IABAC accuracy with increasing IB-blocks length is that in places where the length of ancestral haplotype mosaics is small, large IB-blocks ignore them. In fact, in determining the points of ancestry change, the IB-blocks defined with L_w length play a role in smoothing ancestry change of the SNPs. The longer the L_w length is, the greater smoothing would be, and with the smaller L_w length, less smoothing will be observed. The ancestry inference of IABAC, which is based on ancestral change points, is preciser than the fixed window method where the length of the haplotype blocks is constant.

In determining the points of ancestry change, the correlation between the SNPs is considered. Due to the fact that the alleles within the ancestral population are not independent from each other in dense SNPs (background LD) (4), and also in examining the possibility of ancestry change between two adjacent SNPs using the distance between the alleles of an admixed individual and the mean of the alleles of ancestral populations, the distance of alleles in an SNP alone cannot have much information, so several SNPs must be considered together. As the results have shown, with the increase of the number of SNPs in each block, the accuracy of IABAC increases. In this case, the adjacent SNPs play a role in determining SNPs ancestral information and influence it. IB-blocks are sliding and move between all SNPs, so adjacent SNPs will be in the same IB-block at least once and will play a role in determining ancestral information of each other. As mentioned, the genomes of admixed individuals are a mosaic of different ancestral populations, so close SNPs are more likely to be from the same ancestors than far SNPs. For this reason, in determining the IB-block, in addition to the desired SNP, its nearest neighbors are considered.

In the stated algorithm to investigate the existence of ancestry change between two SNPs, in addition to the left and right IB-blocks of the desired SNPs (red IB-blocks in Figure 3a), the total IB-blocks (blue IB-blocks in Figure 3c) are also considered. The distance between different ancestral populations and admixed individual in the left, right and total IB-blocks is calculated. In fact, IABAC investigates the association between the ancestry of two IB-blocks and the sameness of their ancestral population. In other words, the proposed way for considering the association between the ancestry of IB-blocks is defined based on the statistic D explained for LD, which examines the probability of two alleles occurring at two chromosomal sites and calculates their correlation (40, 41). The value of D between alleles A and B at two chromosomal sites is presented by Equation 10.

$$D_{AB} = P_{AB} - P_A P_B \tag{10}$$

In this equation, D_{AB} represents the value of LD between alleles A and B, P_A denotes the frequencies of alleles A in the population, P_B denotes the frequencies of alleles B in the population, and P_{AB} is the frequencies of alleles A and B together (AB) in the population. Bigger D means more dependency between alleles and smaller D means more independency between alleles. In Equation 10, the frequencies of occurrence of each allele and the sum of alleles are considered, while in the IABAC, the distance of each IB-block and the sum of IB-blocks is considered.

Among the four classification methods mentioned, random forest method more accurately classifies haplotype blocks. With increasing admixture times, the ancestry inference accuracy decreases in all methods. However, the accuracy of IABAC-RF is better than other methods of ancestry inference.

5. CONCLUSION

IABAC infers ancestry by identifying the approximate boundaries of the ancestry change. To identify the boundaries of ancestry changes, IABAC uses distance between ancestral populations by considering the information of the neighbors in SNPs. After identifying boundaries of ancestry change, the ancestry between boundaries is determined. The important features of the IABAC compared to the former methods is that IABAC does not require many statistical or biological parameters. The only input parameter is the length of the IB-block. The results were shown that IABAC with IB-block length of 100-300 SNPs had the highest accuracy of ancestry inference and with increasing admixture times, IABAC with random forest classifier was more accurate that other methods for the ancestry inference. When IABAC and fixed window use the same classifiers, IABAC are more precise than fixed window for ancestry inference. In the present study, we studied haplotype data. We used SHAPEIT to convert genotype to haplotype. Future work will include adding a phasing step to IABAC. With the addition of the phasing step, if we have genotype data, phasing step converts genotype into haplotypes and then IABAC infers the ancestry of haplotypes. Sometimes we need to infer the ancestry of genotype data (similar to EILA). The algorithm used in IABAC, with minor changes, can infer the ancestry of genotype data, which can be done as future work.

6. REFERENCES

1. Cavalli-Sforza LL, Feldman MW. The application of molecular genetic approaches to the study of human evolution. Nature genetics. 2003;33(Suppl 3):266-75. https://doi.org/10.1038/ng1113
2. Yang JJ, Cheng C, Devidas M, Cao X, Fan Y, Campana D, et al. Ancestry and pharmacogenomics of relapse in acute lymphoblastic leukemia. Nature genetics. 2011;43(3):237-41. https://doi.org/10.1038/ng.763
3. Koehl AJ. Estimating ancestry and genetic diversity in admixed populations: The University of New Mexico; 2016.
4. Geza E, Mugo J, Mulder NJ, Wonkam A, Chimusa ER, Mazandu GK. A comprehensive survey of models for dissecting local ancestry deconvolution in human genome. Briefings in bioinformatics. 2019;20(5):1709-24. https://doi.org/10.1093/bib/bby044
5. Price AL, Tandon A, Patterson N, Barnes KC, Rafaels N, Ruczinski I, et al. Sensitive detection of chromosomal segments of distinct ancestry in admixed populations. PLoS genetics. 2009;5(6):e1000519. https://doi.org/10.1371/journal.pgen.1000519
6. Gravel S. Population genetics models of local ancestry. Genetics. 2012;191(2):607-19.
7. Hu Y, Willer C, Zhan X, Kang HM, Abecasis GR. Accurate local-ancestry inference in exome-sequenced admixed individuals via off-target sequence reads. The American Journal of Human Genetics. 2013;93(5):891-9. https://doi.org/10.1016/j.ajhg.2013.10.008
8. Ma Y, Zhao J, Wong J-S, Ma L, Li W, Fu G, et al. Accurate inference of local phased ancestry of modern admixed populations. Scientific reports. 2014;4(1):5800. https://doi.org/10.1038/srep05800
9. Durand EY, Do CB, Mountain JL, Macpherson JM. Ancestry composition: a novel, efficient pipeline for ancestry deconvolution. biorxiv. 2014:010512. https://doi.org/10.1101/010512
10. Khayatzadeh N, Mészáros G, Gredler B, Schnyder U, Curik I, Sölkner J. Prediction of global and local Simmental and Red Holstein Friesian admixture levels in Swiss Fleckvieh cattle. Poljoprivreda. 2015;21(1 SUPPLEMENT):63-7. https://doi.org/10.18047/poljo.21.1.sup.14
11. Alizadeh F, Jazayeriy H, Jazayeri O, Vafaee F, editors. SMIA: a simple way for inference of admixed population ancestors. 2020 10th International Conference on Computer and Knowledge Engineering (ICCKE); 2020: IEEE. https://doi.org/10.1109/ICCKE50421.2020.9303686
12. Pool JE, Nielsen R. Inference of historical changes in migration rate from the lengths of migrant tracts. Genetics. 2009;181(2):711-9. https://doi.org/10.1534/genetics.108.098095
13. Pasaniuc B, Zaitlen N, Lettre G, Chen GK, Tandon A, Kao WL, et al. Enhanced statistical tests for GWAS in admixed populations: assessment using African Americans from CARe and a Breast Cancer Consortium. PLoS genetics. 2011;7(4):e1001371. https://doi.org/10.1371/journal.pgen.1001371
14. Wang X, Zhu X, Qin H, Cooper RS, Ewens WJ, Li C, et al. Adjustment for local ancestry in genetic association analysis of admixed populations. Bioinformatics. 2011;27(5):670-7. https://doi.org/10.1093/bioinformatics/btq709
15. Omberg L, Salit J, Hackett N, Fuller J, Matthew R, Chouchane L, et al. Inferring genome-wide patterns of admixture in Qataris using fifty-five ancestral populations. BMC genetics. 2012;13:1-10. https://doi.org/10.1186/1471-2156-13-49
16. Maples BK, Gravel S, Kenny EE, Bustamante CD. RFMix: a discriminative modeling approach for rapid and robust local-ancestry inference. The American Journal of Human Genetics. 2013;93(2):278-88. http://dx.doi.org/10.1016/j.ajhg.2013.06.020
17. Sankararaman S, Sridhar S, Kimmel G, Halperin E. Estimating local ancestry in admixed populations. The American Journal of Human Genetics. 2008;82(2):290-303. https://doi.org/10.1016/j.ajhg.2007.09.022
18. Alizadeh F, Jazayeriy H, Jazayeri O, Vafaee F. AICRF: Ancestry Inference of Admixed Population with Deep Conditional Random Field. Journal of Genetics. accepted for puplication, 2023. 10.1007/s12041-023-01445-7
19. Paşaniuc B, Sankararaman S, Kimmel G, Halperin E. Inference of locus-specific ancestry in closely related populations. Bioinformatics. 2009;25(12):i213-i21. https://doi.org/10.1093/bioinformatics/btp197
20. Brisbin A, Bryc K, Byrnes J, Zakharia F, Omberg L, Degenhardt J, et al. PCAdmix: principal components-based assignment of ancestry along each chromosome in individuals with admixed ancestry from two or more populations. Human biology. 2012;84(4):343. https://doi.org/10.3378%2F027.084.0401

21. Salter-Townshend M, Myers S. Fine-scale inference of ancestry segments without prior knowledge of admixing groups. Genetics. 2019;212(3):869-89. https://doi.org/10.1534/genetics.119.302139
22. Guan Y. Detecting structure of haplotypes and local ancestry. Genetics. 2014;196(3):625-42. https://doi.org/10.1534/genetics.113.160697
23. Kumar A, Montserrat DM, Bustamante C, Ioannidis A. Xgmix: Local-ancestry inference with stacked xgboost. BioRxiv. 2020:2020.04. 21.053876. https://doi.org/10.1101/2020.04.21.053876
24. Montserrat DM, Bustamante C, Ioannidis A, editors. Lai-net: Local-ancestry inference with neural networks. ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP); 2020: IEEE. https://doi.org/10.1109/ICASSP40776.2020.9053662
25. Oriol Sabat B, Mas Montserrat D, Giro-i-Nieto X, Ioannidis AG. SALAI-Net: species-agnostic local ancestry inference network. Bioinformatics. 2022;38(Supplement_2):ii27-ii33. https://doi.org/10.1093/bioinformatics/btac464
26. Wang Y, Song S, Schraiber JG, Sedghifar A, Byrnes JK, Turissini DA, et al. Ancestry inference using reference labeled clusters of haplotypes. BMC bioinformatics. 2021;22(1):1-14. https://doi.org/10.1186/s12859-021-04350-x
27. Browning SR, Browning BL. Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering. The American Journal of Human Genetics. 2007;81(5):1084-97. https://doi.org/10.1086/521987
28. Yang JJ, Li J, Buu A, Williams LK. Efficient inference of local ancestry. Bioinformatics. 2013;29(21):2750-6. https://doi.org/10.1093/bioinformatics/btt488
29. Gaurav K, Kumar A, Singh P, Kumari A, Kasar M, Suryawanshi T. Human Disease Prediction using Machine Learning Techniques and Real-life Parameters. International Journal of Engineering. 2023;36(6):1092-8. https://doi.org/10.5829/ije.2023.36.06c.07
30. Hamidi H, Qaribpour F. An efficient predictive model for probability of genetic diseases transmission using a combined model. International Journal of Engineering. 2017;30(8):1152-9. 10.5829/ije.2017.30.08b.06
31. Kumar S, Sahoo G. A random forest classifier based on genetic algorithm for cardiovascular diseases diagnosis. International Journal of Engineering, Transactions B: Applications. 2017;30(11):1723-9. 10.5829/ije.2017.30.11b.13
32. Zamani F, Mohammadjani A. A Multiple Kernel Learning based Model with Clustered Features for Cancer Stage Detection using Gene Datasets. International Journal of Engineering, Transactions B: Applications. 2023. https://doi.org/10.5829/ije.2023.36.11b.08
33. Shedthi B S, Shetty V, Chadaga R, Bhat R, Bangera P, Kini K P. Implementation of Chatbot that Predicts an Illness Dynamically using Machine Learning Techniques. International Journal of Engineering. 2023. IJE Article in press
34. Anbananthen KSM, Busst MBMA, Kannan R, Kannan S. A Comparative Performance Analysis of Hybrid and Classical Machine Learning Method in Predicting Diabetes. Emerging Science Journal. 2022;7(1):102-15. https://doi.org/10.28991/ESJ-2023-07-01-08
35. Muthaiyah S, Singh VA, Zaw TOK, Anbananthen KS, Park B, Kim MJ. A Binary Survivability Prediction Classification Model towards Understanding of Osteosarcoma Prognosis. Emerging Science Journal. 2023;7(4):1294-314. https://doi.org/10.28991/ESJ-2023-07-04-018
36. Frazer KA, Ballinger DG, Cox DR, Hinds DA, Stuve LL, Gibbs RA, et al. A second generation human haplotype map of over 3.1 million SNPs. Nature. 2007;449(7164):851-61. https://doi.org/10.1038%2Fnature06258
37. Delaneau O, Coulonges C, Zagury J-F. Shape-IT: new rapid and accurate algorithm for haplotype inference. BMC bioinformatics. 2008;9(1):1-14. https://doi.org/10.1186%2F1471-2105-9-540
38. Geza E, Mulder NJ, Chimusa ER, Mazandu GK. FRANC: a unified framework for multi-way local ancestry deconvolution with high density SNP data. Briefings in bioinformatics. 2020;21(5):1837-45. https://doi.org/10.1093/bib/bbz117
39. Myers S, Bottolo L, Freeman C, McVean G, Donnelly P. A fine-scale map of recombination rates and hotspots across the human genome. Science. 2005;310(5746):321-4. https://doi.org/10.1126/science.1117196
40. Slatkin M. Linkage disequilibrium—understanding the evolutionary past and mapping the medical future. Nature Reviews Genetics. 2008;9(6):477-85. https://doi.org/10.1038/nrg2361
41. Smith RD. The nonlinear structure of linkage disequilibrium. Theoretical Population Biology. 2020;134:160-70. https://doi.org/10.1016/j.tpb.2020.02.005

Persian Abstract

چکیده

ترکیب ژنوم یک پدیده رایج در جمعیت های انسانی است که نتیجه آمیزش افراد از دو یا چند جمعیت مستقل است. این می تواند منجر به تشکیل DNA جدیدی شود که هر بخش آن از جمعیت اجدادی متفاوت منشا می گیرد. روش‌های استنباط تبار محلی برای شناسایی تبار هر بخش از ژنوم استفاده می‌شوند که می‌تواند بینشی در مورد تاریخچه جمعیت‌ها ارائه دهد. بسیاری از روش‌های استنباط تبار محلی (LAI) نیاز به تعیین پارامترهای مختلفی دارند که ممکن است بدست آوردن آنها دشوار بوده و استفاده از روش‌های LAI را با مشکل مواجه کند.در این مقاله، ما یک روش جدید برای شناسایی مرزهای تقریبی تغییر تبار (IABAC) در هاپلوتیپ‌های مخلوط و سپس تعیین تبار بین مرزها ارائه می‌دهیم. بر خلاف بسیاری از روش های LAI، روش ما به پارامترهای آماری یا بیولوژیکی زیادی متکی نیست، بنابراین در برابر تغییرات الگوهای اختلاط قوی‌تر است .ما روش خود را بر روی داده های انسان ارزیابی می کنیم و نشان می دهیم که از روش های موجود تشخیص تبار دقیق تر است. نتایج ما نشان می‌دهد که IABAC یک روش جدید امیدوارکننده برای شناسایی مرزهای تغییر تبار در هاپلوتیپ‌های مخلوط است. این روش می تواند برای مطالعه تاریخچه اختلاط در جمعیت های انسانی و شناسایی واریانت های ژنتیکی که با جمعیت های اجدادی مختلف مرتبط هستند مورد استفاده قرار گیرد.

International Journal of Engineering

Journal Homepage: www.ije.ir

Improving Seismic Vulnerability of Irregular Reinforced Concrete Moment-Resisting Frames using Shear Walls

M. Mouhine*[a], M. Derife[b], S. Aboumdian[c], E. Hilali[d]

[a] *ASE Laboratory, National School of Applied Sciences, Ibn Tofail University, Morocco*
[b] *GEMS Laboratory, National School of Applied Sciences, Ibn Zohr University, Morocco*
[c] *LSIB, Faculty of Sciences and Techniques Mohammadia, Hassan 2 University, Morocco*
[d] *MMGC Laboratory, National School of Applied Sciences, Ibn Zohr University, Morocco*

PAPER INFO

Paper history:
Received 18 September 2023
Received in revised form 03 October 2023
Accepted 04 October 2023

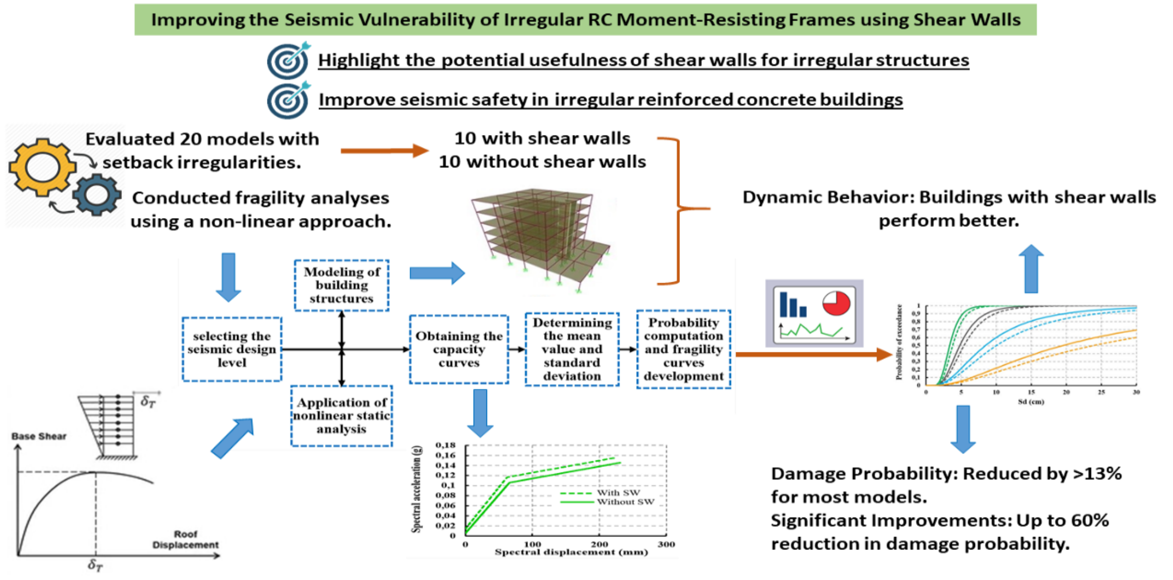
Keywords:
Reinforced Concrete Buildings
Seismic Vulnerability
Vertical Geometric Irregularity
Fragility Curves
Non-linear Analysis

ABSTRACT

Vertical geometric irregular reinforced concrete (RC) buildings are widely used in structural engineering due to their aesthetic appearance and functional characteristics. Indeed, improving their reliability and seismic performance is of crucial interest and has even become a necessity. This research study underlines the importance of using shear walls (SW) as a fundamental means of reinforcement for this type of structure. Twenty models, including ten with SW and ten without SW, of mid-rise buildings with setback irregularity were considered for this purpose, and fragility analyses were carried out, using a non-linear procedure, to highlight the potential usefulness of shear walls for irregular structures. The results of this work clearly indicate that the dynamic behavior and response of buildings have been improved by the use of shear walls. The fragility study reveals that for some cases the damage probability is reduced, with the difference exceeding 13% for the majority of models, and for some cases the differences are highly significant, ranging from 30% to 60%. This shows the benefits of incorporating shear walls into the design phase of irregular buildings.

***doi**: 10.5829/ije.2024.37.02b.17*

Graphical Abstract



Improving the Seismic Vulnerability of Irregular RC Moment-Resisting Frames using Shear Walls

Highlight the potential usefulness of shear walls for irregular structures

Improve seismic safety in irregular reinforced concrete buildings

Evaluated 20 models with setback irregularities.

Conducted fragility analyses using a non-linear approach.

10 with shear walls
10 without shear walls

Dynamic Behavior: Buildings with shear walls perform better.

Damage Probability: Reduced by >13% for most models.
Significant Improvements: Up to 60% reduction in damage probability.

---

*Corresponding Author Email: medmouhine@gmail.com (M. Mouhine)

Please cite this article as: Mouhine M, Derife M, Aboumdian S, Hilali E. Improving Seismic Vulnerability of Irregular Reinforced Concrete Moment-Resisting Frames using Shear Walls. International Journal of Engineering, Transactions B: Applications. 2024;37(02):425-38.

1. INTRODUCTION

Recently, in the field of structural engineering, the use of irregular building structures has become increasingly common, due to their functional characteristics and aesthetic appearance (1-3). Despite scientific and technical progress in the fields of geotechnics and earthquake engineering, this type of structure has shown a mysterious behavior when subjected to seismic shocks. The earthquakes that have shaken many countries around the world have confirmed that the majority of economic and human losses due to seismic activity result mainly from the defective behavior of building structures.

Structural vulnerability assessment is a fundamental element of modern performance-based seismic design and assessment procedures. Major advances in the development and implementation of fragility functions have taken place over the last three decades. Structural seismic fragility refers to the probability of a structure experiencing failure or reaching a specific failure state during an earthquake (4-7). Understanding the structural seismic vulnerability of structures is crucial for designing and constructing resilient buildings that can withstand seismic events. It helps engineers identify potential weaknesses and implement appropriate measures to mitigate the risk of failure, ensure occupant safety, and minimize damage during earthquakes. Methods for assessing seismic fragility curves differ in complexity, accuracy, and purpose. The choice of method for assessing seismic vulnerability will depend on the quality and quantity of data available, and on the objective, which may be to estimate the seismic vulnerability of a single building or a group of buildings (8). This extensive work reflects, on the one hand, the importance attached by researchers to estimating the damage to different systems and, on the other, the multiplicity of approaches developed. Among the methods listed, the large majority are analytical. Empirical methods also featured prominently in these studies.

In this context, empirical seismic vulnerability has been widely applied worldwide to assess seismic vulnerability and risk for different structural systems. Li and Gardoni (9) have studied the impact of multidirectional seismic sequences, updated the instrument intensity calculation model, and proposed a quantitative method taking into account hybrid intensity measurements to assess the vulnerability of building groups. Also, research studies conducted by Li and his coworkers (5-7, 10-12) have shown the importance and interest of assessing the vulnerability and resilience of structures on the basis of empirical studies.

Ultimately, it can be contended that the choice of the most appropriate procedure depends on the resources available for data collection, the computational expertise available, and finally the scale and objective of the study. In this research work, to estimate the seismic vulnerability of irregular building structures, the analytical approach was chosen to determine the seismic response of the structure through a non-linear static analysis.

The seismic reliability of buildings under seismic actions is greatly affected and the vulnerability of the building to damage caused by ground motions becomes more significant, particularly for building systems with structural irregularities. Therefore, it is interesting to focus on the response and behavior of these structures in seismic conditions. This need has aroused the interest of scientists in this field, and for decades there has been a steady stream of research aimed at developing seismic design methodologies to better understand the dynamic behavior of irregular buildings and to reduce their consequences. Over the past few years, many researchers have studied the influence of setback irregularity on the dynamic response of RC building. They have used the concept of fragility analysis to carry out both static and dynamic analyses. Among these studies, Ruggieri (13), Praveen and Gopikrishna (14) Azad et al. (15), Shojaei and Behnam (16), Ruggieri et al. (17, 18) focused on investigating the local and global performance of setback RC buildings designed according to different international standards. According to the researchers, the geometric irregularity of structures obviously influences the seismic response of building structures, which has a remarkable effect on the probability of damage of buildings. Likewise, research works conducted by Men et al. (19), Kassem et al. (20), Nazri et al. (21), Kumar et al. (22), Ayub et al. (23), Mouhine and Hilali (24, 25), El Janous and El Ghouloubzouri (26), Hashim and Ali (27) show that the presence of irregularities in the structural configuration affect significantly the performance and the dynamic response of reinforced concrete buildings during the seismic excitation. In construction engineering, the use of shear walls is a highly functional alternative for improving the resistance of structures to lateral and gravity loads. Shear walls are essential for achieving very good performance under extreme load conditions (28). Also, shear wall systems offer greater lateral rigidity to effectively reduce displacement and maintain the structural integrity during earthquake events (29).

This work's main purpose is to investigate the contribution of shear wall-resisting systems to enhance the seismic reliability of vertical geometric irregular buildings. Storey displacement, an important indicator of structural performance, is used to assess how this solution works to minimize the storey displacement and limit building drift ratio. Moreover, structural capacity and fragility curves will be investigated to determine how successfully the shear wall system dissipates energy and resists to seismic consequences.

The outcomes of this research work provide substantial practical benefits for setback building structures design in seismic regions. This research explores very useful perspectives for the selection and improvement of building systems to assure the structural integrity and resilience of structures exposed to dynamic loads by contrasting the dynamic behavior of irregular buildings with and without shear wall systems. Understanding the benefits and drawbacks of these systems will also contribute to advancing earthquake design procedures and facilitating informed decision-making.

In summary, this research study highlights how crucial it is to consider the importance of using shear walls system to improve the dynamic behavior of irregular RC structures. Additionally, the study intends to expand knowledge of the role and performance of shear walls for irregular buildings by conducting a rigorous analysis. As such, researchers and practitioners involved in structural design and rehabilitation are likely to benefit considerably from the results of this research work.

2. MODELING AND METHODS

2. 1. Building Structures Description and Modeling

In the present study, twenty models of six-storey moment-resisting reinforced concrete frames are selected, including 10 models with shear walls, as illustrated in Figure 1. These configurations represent mid-rise residential building structures. The selected models are located in Agadir city in Morocco which is qualified as a high seismic risk region according to the Moroccan standard. The studied buildings have a total height of 18 meters, with six floors, and each floor height of 3 meters. These building models have a plan area of 300 meters square and an identical plan dimension of 20m × 15m with five bays in the longitudinal direction and three bays in the transverse direction. The dimension of the beams and columns of the structure has been taken at 200mm × 400mm and 400mm × 400mm, respectively with a reinforcement bars of 3T10 + 3T14 for beams and 8T16 for columns. The shear walls have a thickness of 200mm with T12 reinforcement bars in both directions. The slabs are 150mm thick, supposed to be rigid, and support their self-mass as well as additional loads of 2.5 KN/m² for live loads (LL) and 1.5 KN/m² for dead loads (DL). The combination (DL) + 0.2 (LL) is used to take into account the structure’s weight according to RPS2000. In this work, seismic analyses were carried out using a FE program (30). The studied configurations are modeled as 3D models including the modeling of all building components such as slab, beam, column, shear

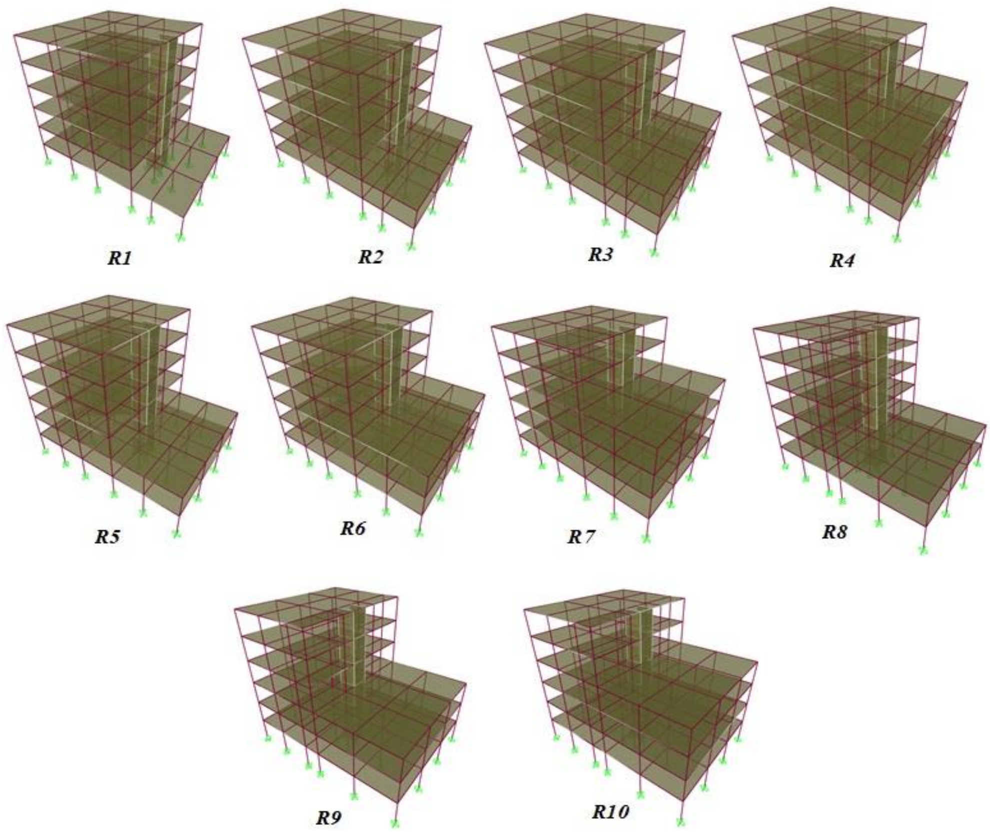


Figure 1. Configurations of building structures with shear walls

wall, and boundary conditions. Strong-column weak-beam concept is used to reduce significantly the probability of collapse (31). Nonlinearities are considered in structural elements by defining plastic hinges in beams and columns (32-34). For the materials used in the building construction, the concrete is class C25/30 with a characteristic compressive strength of 25 MPa. Tensile strength and modulus of elasticity are calculated in accordance with European standard (EC2) (35). The reinforcement bars have a yield strength of 400 MPa and a Young modulus of 210000 MPa. The bars are spaced following the constructive dispositions specified by EC8-2004 (36).

2. 2. Concept of Fragility Curves

The reliability of irregular reinforced concrete buildings under seismic shaking is an interesting topic in the field of earthquake engineering and seismic risk management. The vulnerability of building structures represents their sensitivity to the damage caused by ground motion (37-41). Many researchers used the fragility curves concept to describe the structural damage states occurred in buildings during and after seismic shaking and to assess the probability of damage to the building structures (16, 22, 24, 28). A fragility curve provides an overview of the building's behavior in response to a particular seismic activity. Fragility curves allow to calculate the probability that a structural engineering demand parameter (d) exceeds a specific damage state (ds) based on a parameter defining the seismic intensity chosen as the spectral displacement (Sd) in this study, as given in Figure 2. The curve of fragility is a cumulative lognormal distribution function with logarithmic standard deviation β and mean value γ (42-46). The mean value γ_{dsi} is defined as a function of yield D_y and ultimate D_u displacement of the structure obtained from the capacity curves of buildings. The expressions retained in this research for the calculation of damage state thresholds in accordance to Milutinovic and Trendafiloski (45) are presented in Table 1.

$$P\left[ds_i / Sd\right] = \Phi\left[\frac{1}{\beta_{ds_i}} \ln\left(\frac{Sd}{\gamma_{ds_i}}\right)\right] \quad (1)$$

Φ : normal cumulative distribution function

γ_{dsi} : mean displacement

β_{dsi} : logarithmic standard deviation computation (24, 47)

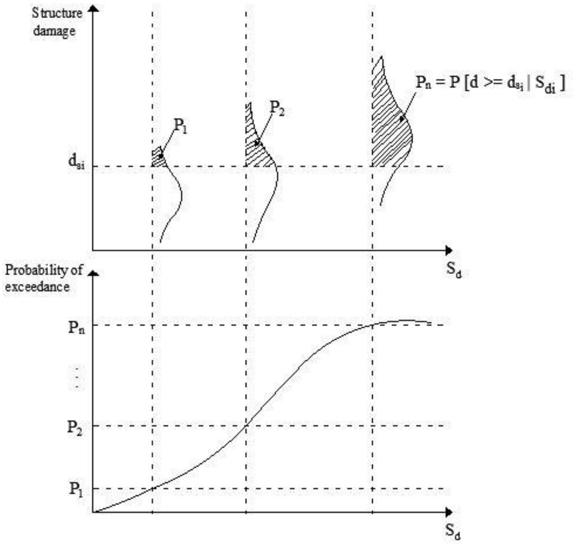


Figure 2. Probability of exceedance

TABLE 1. Mean value formulas [24]

State of Damage	State of Damage Thresholds
Slight damage	$\gamma ds_1 = 0.7 \times D_y$
Moderate damage	$\gamma ds_2 = D_y$
Severe damage	$\gamma ds_3 = D_y + 0.25\left(D_u - D_y\right)$
Complete damage	$\gamma ds_4 = D_u$

Figure 3 depicts the strategy employed in the present work to create fragility curves using nonlinear static (pushover) procedure. In this approach, after selecting the appropriate seismic design level, the building structure models considered are modeled using a FE calculation program (30). Then, these buildings sample are subjected to a seismic load to obtain the capacity and the performance of each structure.

3. RESULTS AND DISCUSSION

3. 1. Storey Displacement Analysis

To emphasize the importance of using shear walls, at the design level, on the structural behavior of irregular buildings, Figure 4 illustrates the results obtained in terms of the displacement, corresponding to the different levels of the building, registered from non-linear static loading. The results clearly indicate that shear walls have a significant influence on structural response in terms of storey displacement. In general, story drift is reduced for the case of buildings with shear walls compared with those without, and this trend is almost the same for all the cases considered in this work. Furthermore, for buildings with vertical geometric irregularity at levels 1 and 2 with a percentage setback of 30%, it is interesting to note that the effect of shear walls is minimal, and the maximum

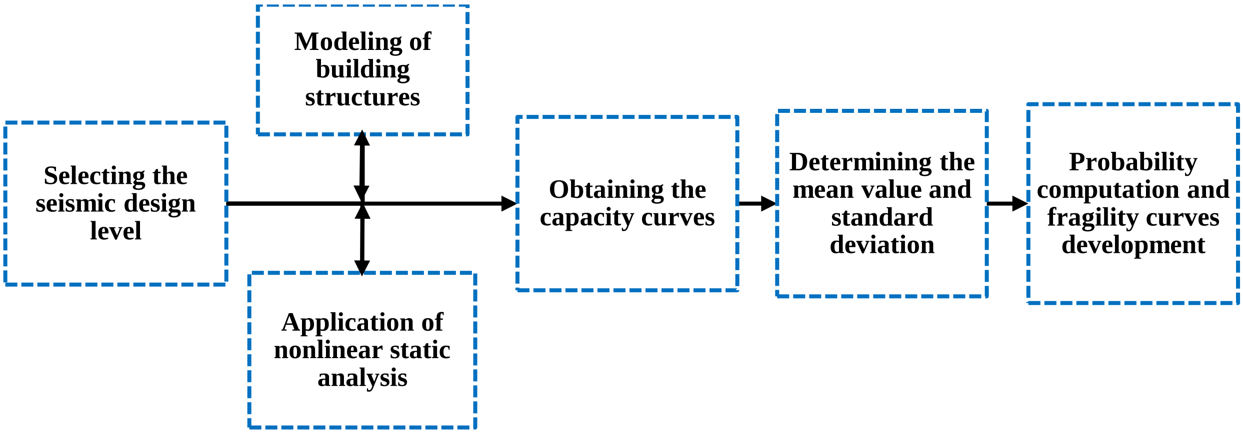
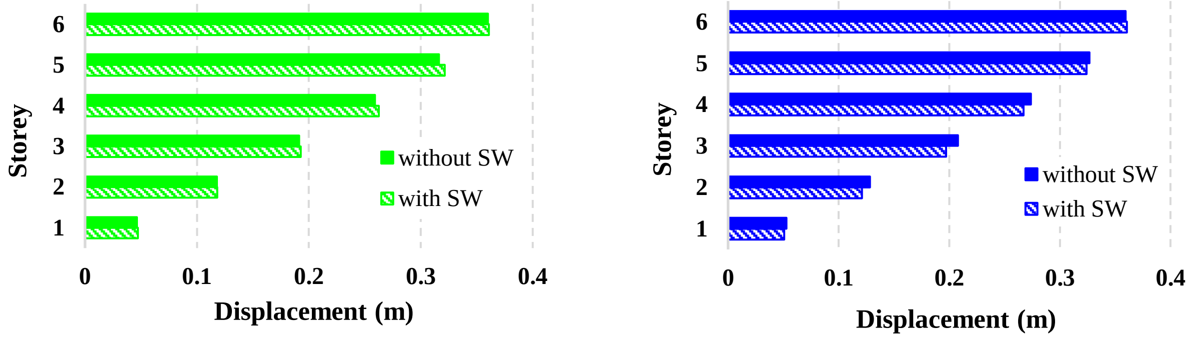
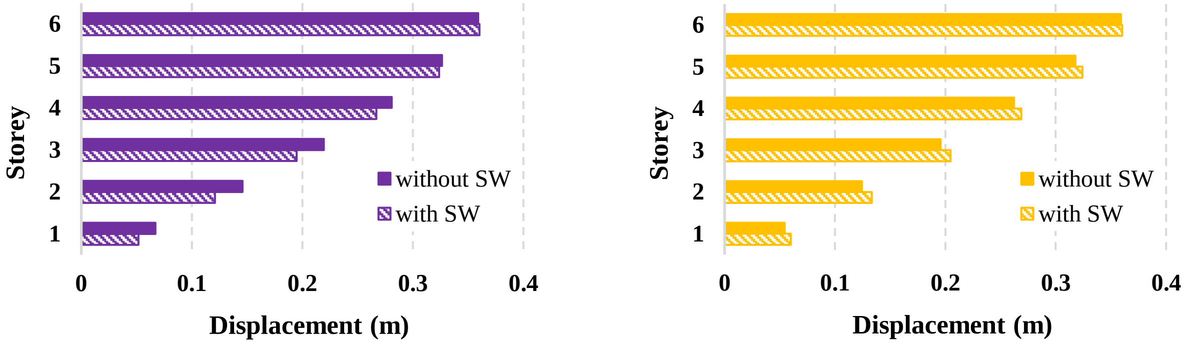


Figure 3. Diagram of fragility curves development strategy



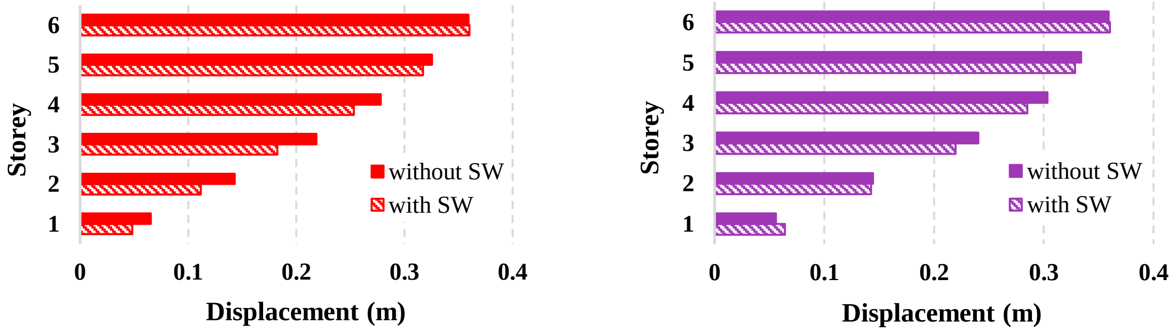
(a) for R1 model

(b) for R2 model



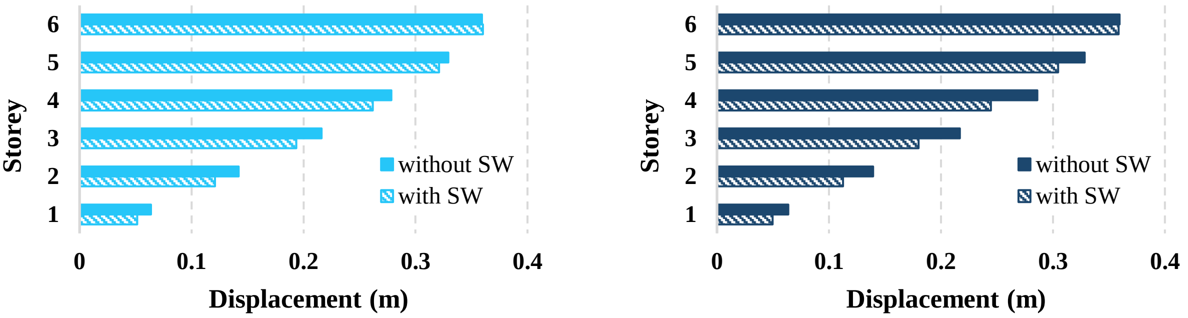
(c) for R3 model

(d) for R4 model



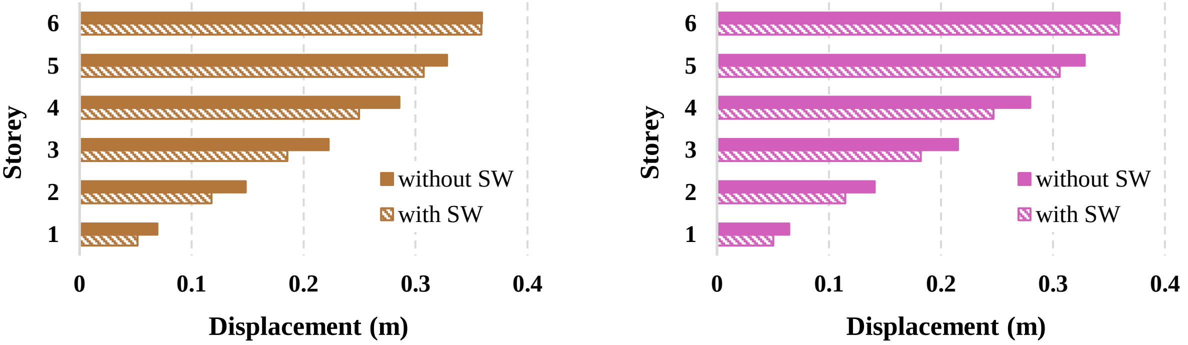
(e) for R5 model

(f) for R6 model



(g) for R7 model

(h) for R8 model



(i) for R9 model

(j) for R10 model

Figure 4. Storey displacement of building structures

deviation does not exceed 2.32% and 3.75% recorded for the case of models R1 and R2, respectively. For the other cases studied, the analysis indicates that the improvement in building response is significant and the influence of shear walls is remarkable, particularly for building models with relatively large setback values, between 40% and 50%, where the deviation can reach values of 15% to 25%, as in the case of R5, R7, R8, R9 and R10 models.

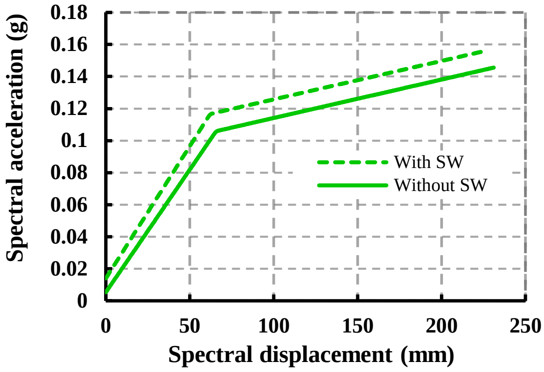
3. 2. Analysis of Building Structure Capacity

The analysis of the dynamic behavior, using a non-linear static approach, of the building structures considered in this work has highlighted the importance and the effect of using shear walls as an effective means of strengthening and improving the response of irregular buildings under seismic actions. Figure 5 shows the capacity curves, in bilinear form, obtained for all the building models considered in this study. It is clear from these figures that the use of shear walls enhances the response of buildings. This consequence can be explained by the fact that shear walls provide stiffness and strength to irregular structures, thus compensating for the stiffness losses experienced in the case of setback structures due to the reduction in structural elements. The results show that for models R8, R9, and R10, the use of shear walls significantly improved the ultimate capacity of the buildings. The differences registered are about 18.36%, 17%, and 14.36%, respectively. This incidence encourages engineers and professionals to build more resilient buildings while retaining the architectural appearance of the structures. For the other models, the difference remains a little less expressive, with a value not exceeding 9.66% observed in the case of the R4 model.

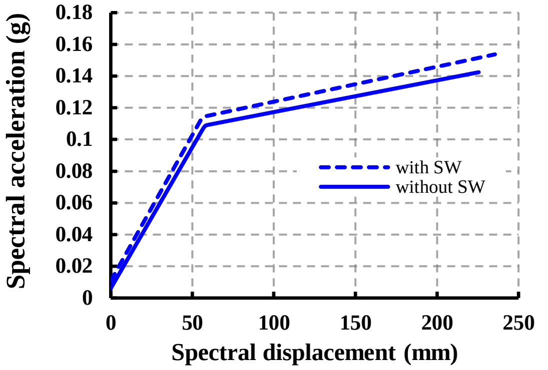
3. 3. Seismic Vulnerability Analysis

In this section, the main aim is to investigate the impact of shear walls on the seismic vulnerability of building structures in the presence of vertical geometric irregularity (setback). For this purpose, seismic reliability analysis of mid-rise buildings is carried out, and fragility curves are generated and compared as shown in Figure 6. The probability of exceeding a particular state of damage is then computed for each building model, and the results are summarized in the histograms shown in Figure 7.

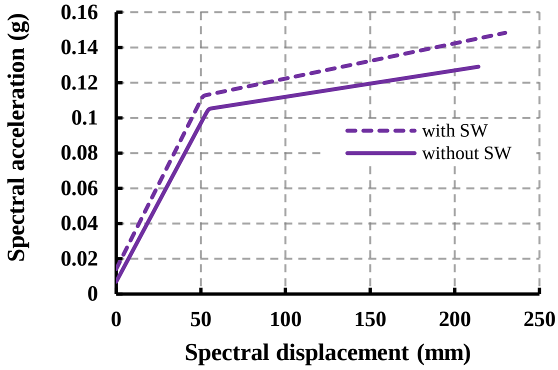
For all the cases studied, it is remarkable that the use of shear walls improves the seismic reliability of buildings, as can be seen in Figure 7. The probability that certain columns and beams near or within the joints suffer cracks due to bending or shear stresses is reduced by 12.19%. This improvement in dynamic behavior of buildings is more pronounced in the case of models R8, R9 and R10, where the percentage rises to 30.72%, 18.07% and 20.66%, respectively. For a moderate damage state, the influence of shear walls on the seismic



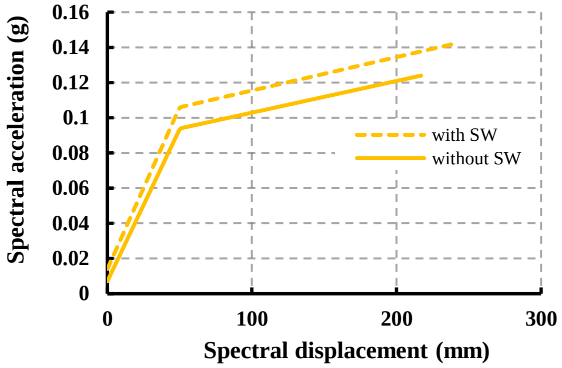
(a) for R1 model



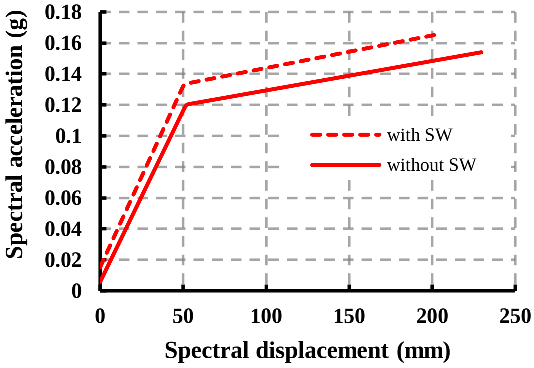
(b) for R2 model



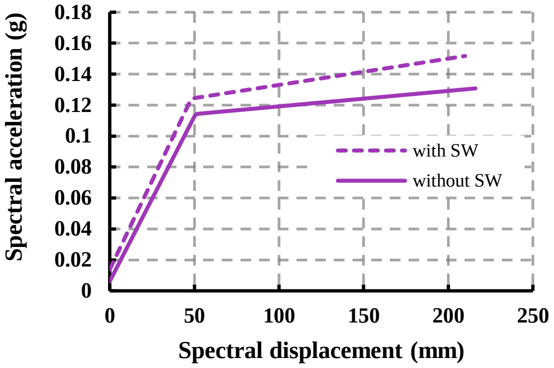
(c) for R3 model



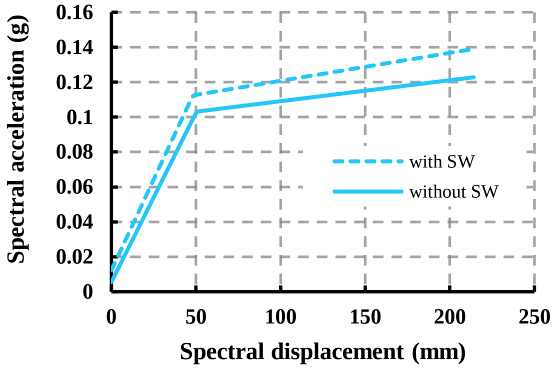
(d) for R4 model



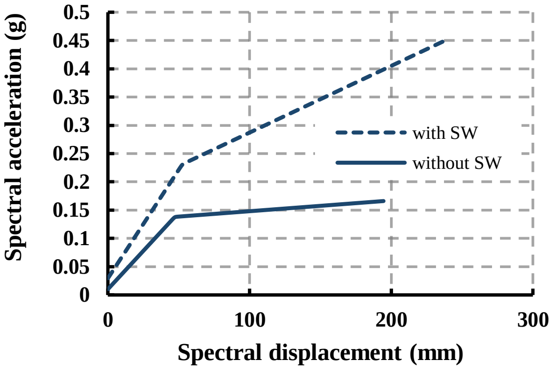
(e) for R5 model



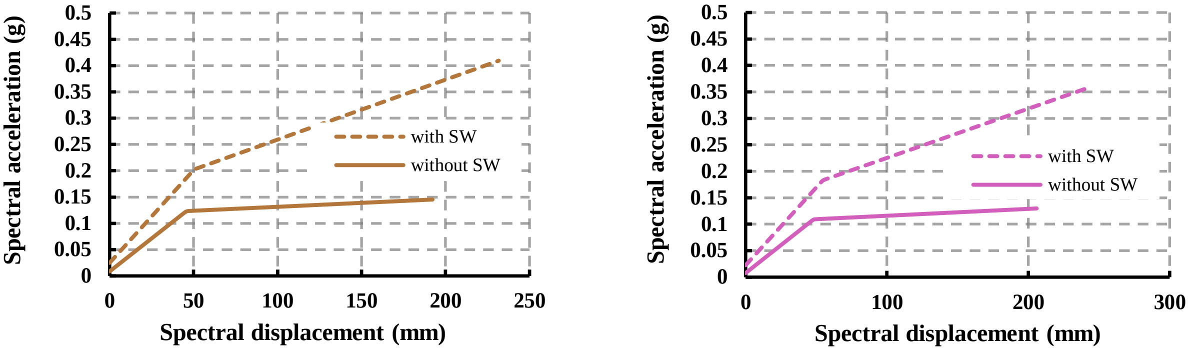
(f) for R6 model



(g) for R7 model

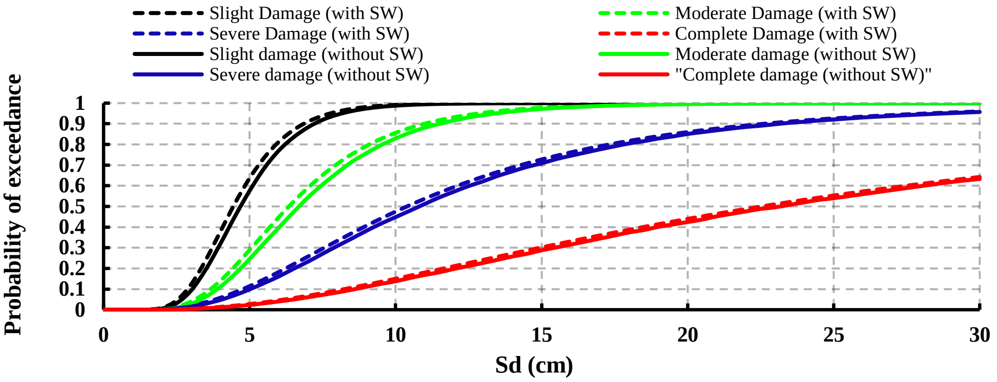


(h) for R8 model

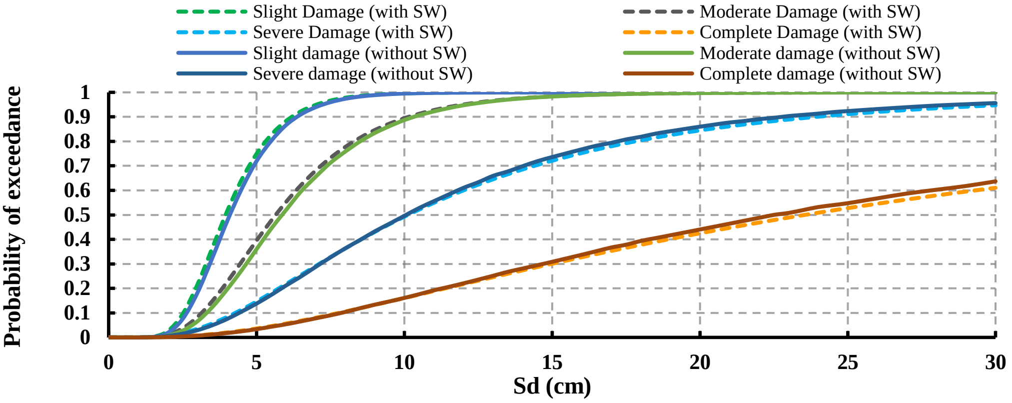


(i) for R9 model (j) for R10 model

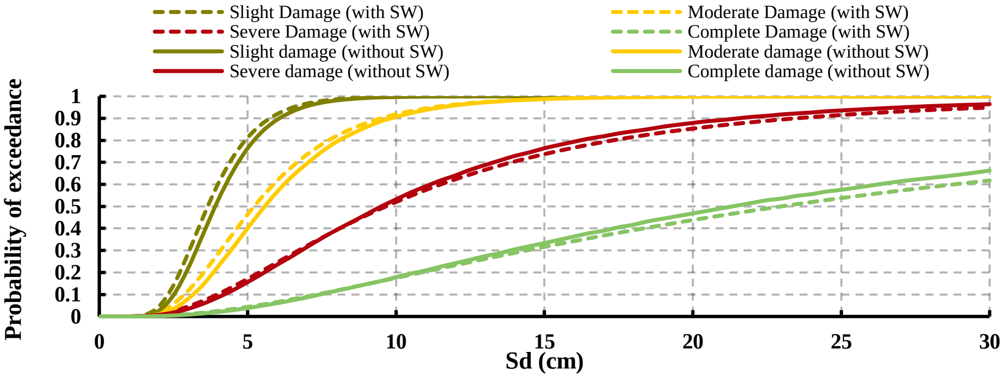
Figure 5. Capacity curves of building structures under consideration



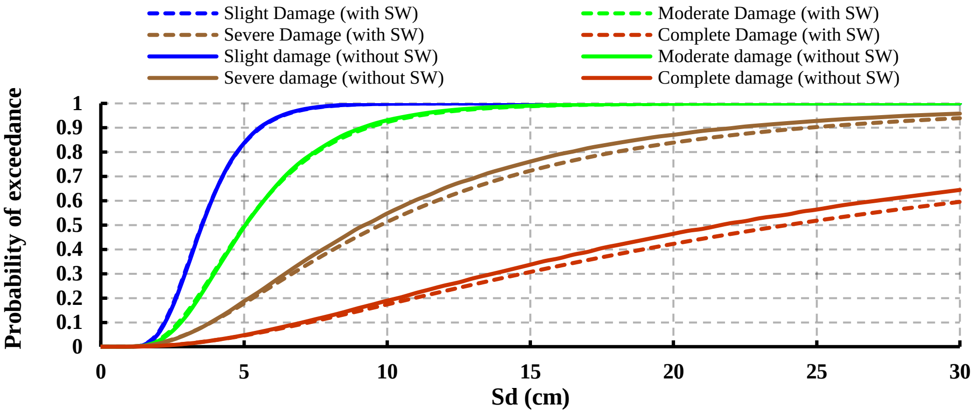
(a) for R1 model



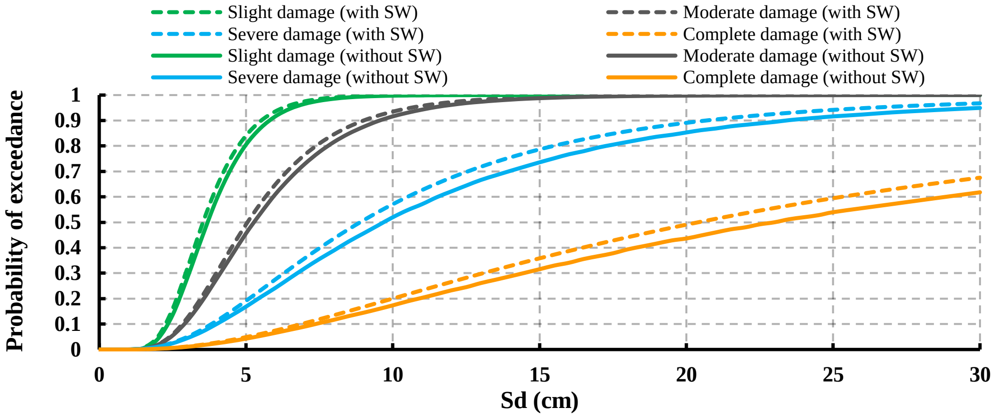
(b) for R2 model



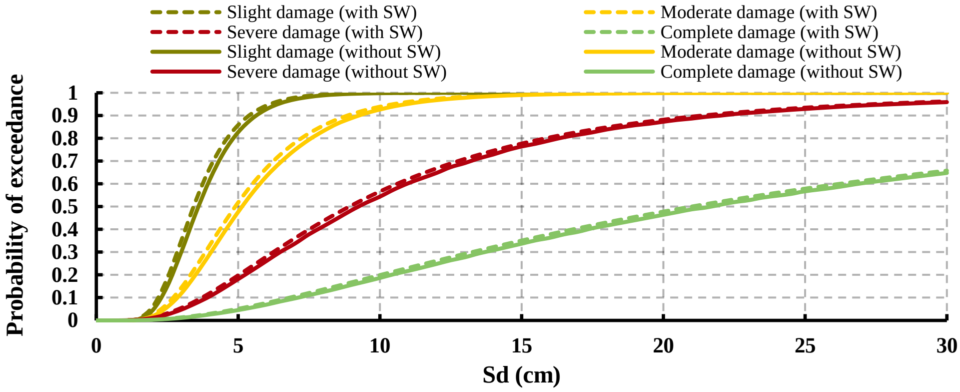
(c) for R3 model



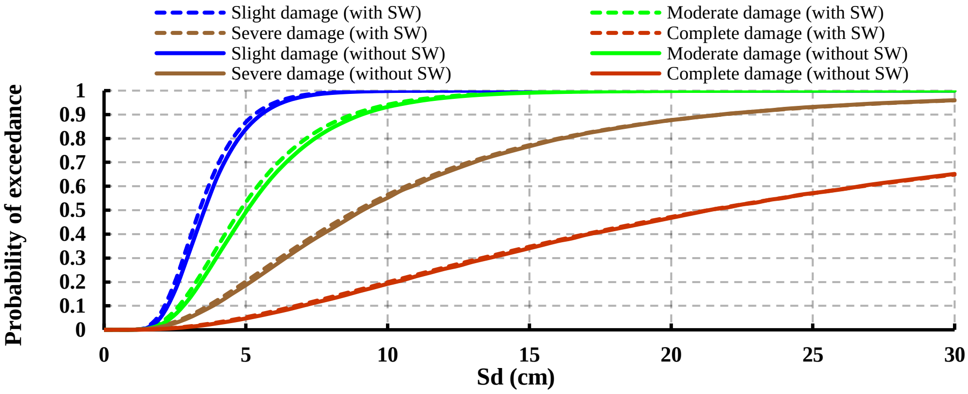
(d) for R4 model



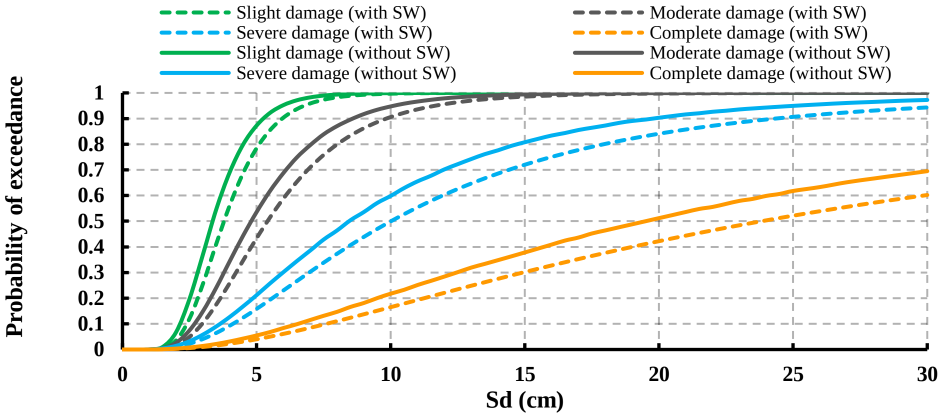
(e) for R5 model



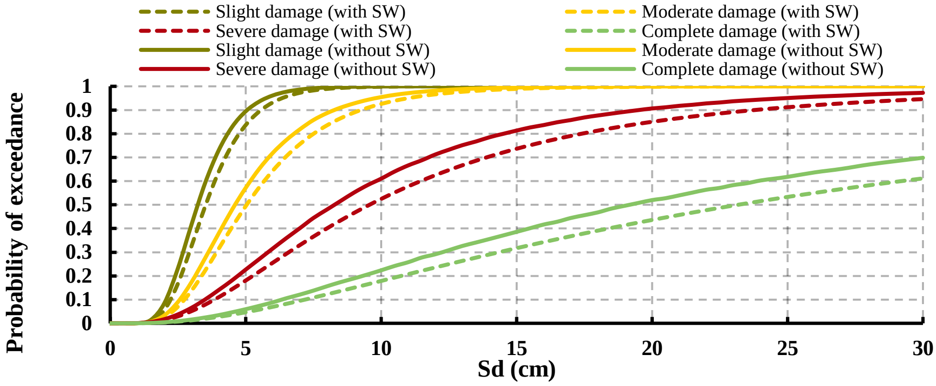
(f) for R6 model



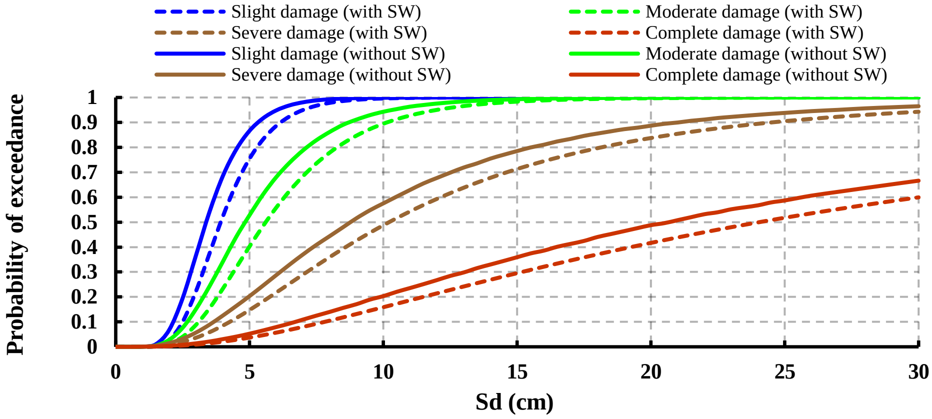
(g) for R7 model



(h) for R8 model



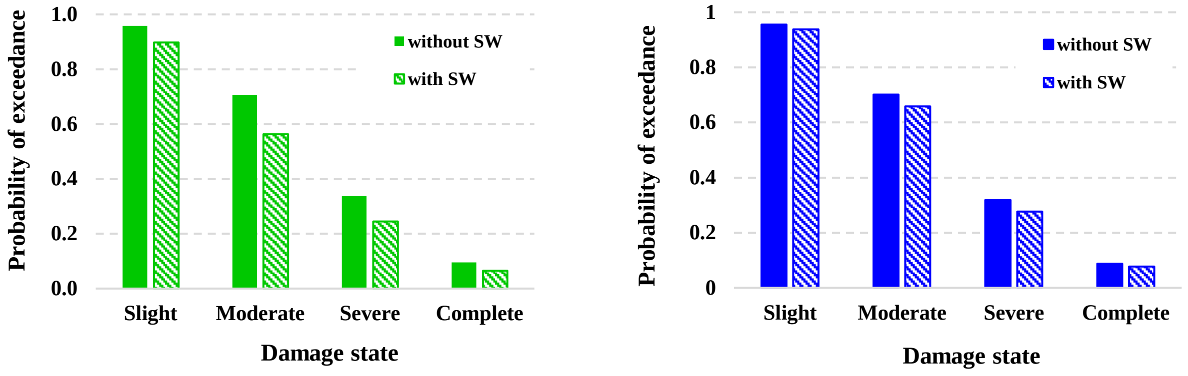
(i) for R9 model



(j) for R10 model

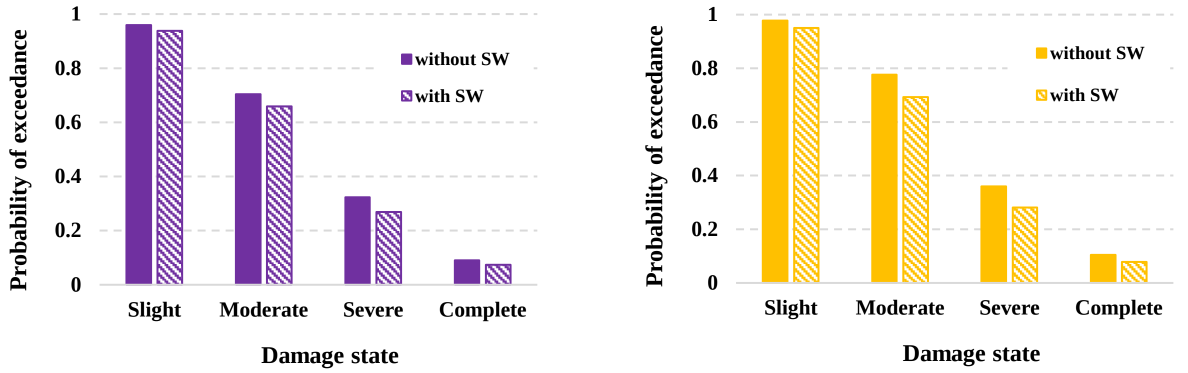
Figure 6. Fragility curves for the studied buildings models

performance of building structure is more apparent. The damage probability is reduced and the difference is about 51.08%, 37.92% and 43.32% for models R8, R9 and R10, respectively. For the other configurations, the difference ranges from 6.02% to 20.26%. Similarly, for a severe damage state, the seismic risk decreases significantly, reaching a value of 58.6% for the R8 model. Considering a complete damage state, the probability of the structure collapsing or presenting a significant collapse risk because of the brittle failure of non-ductile beams and columns elements or loss of frame stability, is reduced to reach maximum values ranging from 63.48% to 54.8% for models R8, R9 and R10. For the other models studied, the differences vary between 13% and 32%. Finally, shear walls contribute, among other things, to the stability of irregular structures by improving their dynamic behavior. The probability of damage is reduced, and the differences can be significant compared to structures without shear walls.



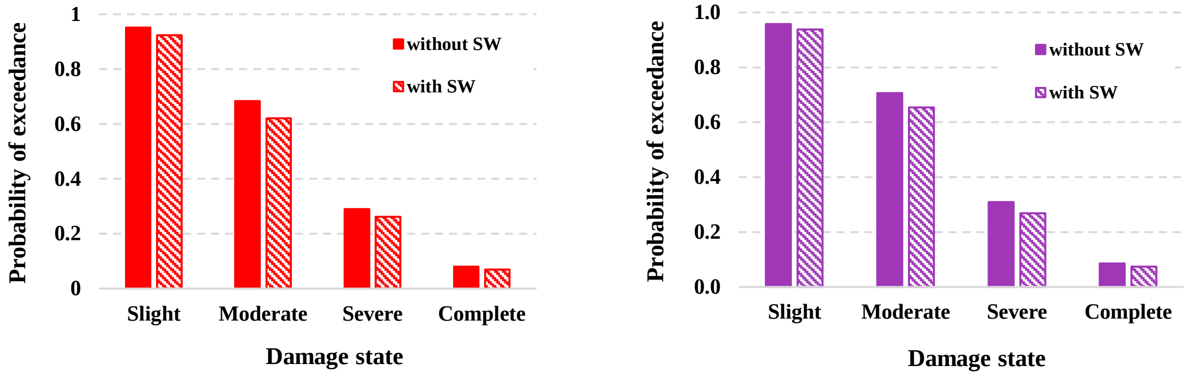
(a) for R1 model

(b) for R2 model



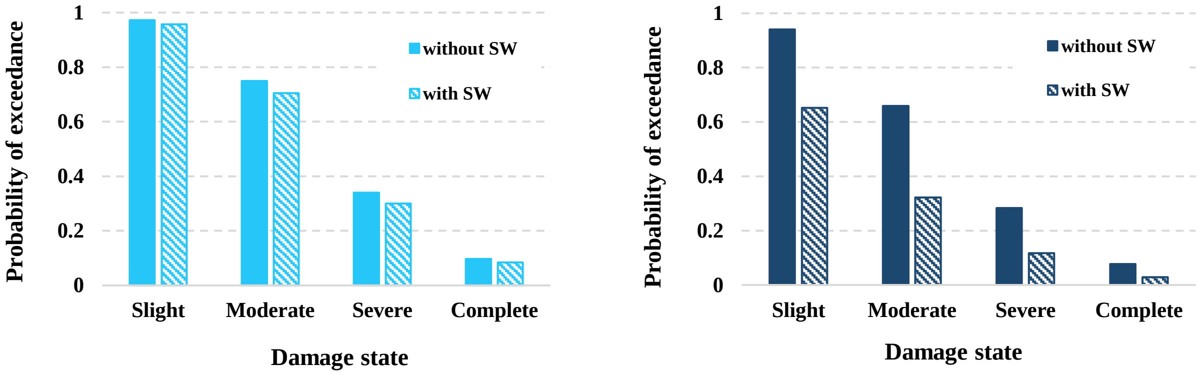
(c) for R3 model

(d) for R4 model



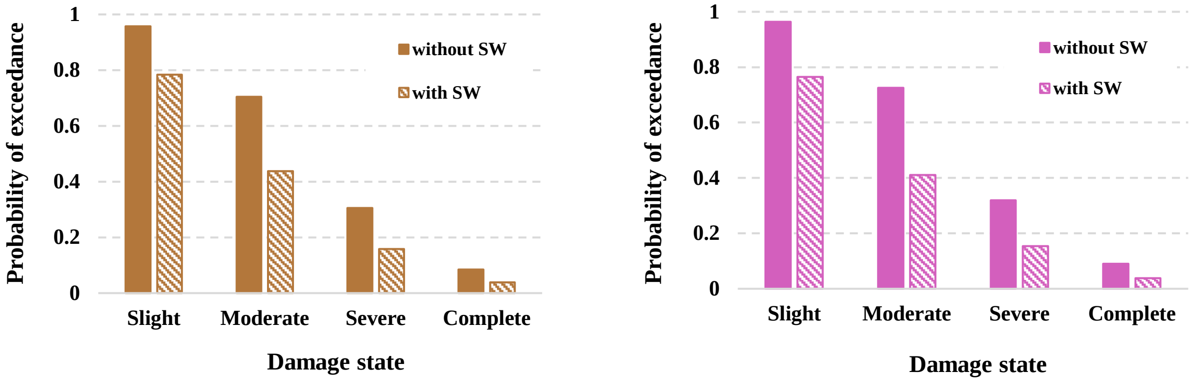
(e) for R5 model

(f) for R6 model



(g) for R7 model

(h) for R8 model



(i) for R9 model (j) for R10 model

Figure 7. Damage probabilities of building structures

4. CONCLUSIONS

This study highlights the importance of using shear walls (SW) as a means of reinforcement to enhance the seismic reliability of irregular RC buildings. The outcomes of this investigation clearly indicate that the capacity of structures is improved when shear walls are used. The difference is about 10% to 18%, which represents a significant deviation that affects the dynamic behavior of these structures. The use of shear walls provides a useful means of compensating losses in stiffness and strength due to the reduction in structural elements. Furthermore, the analysis shows that the vulnerability is highly reduced in the case of building with shear walls, the deviation in terms of fragility is interesting and exceeds 30% for several tested building models. Also, the results indicate that the storey displacement is reduced and the deviation can reach values of 15% to 25% as in the case of R5, R7, R8, R9, and R10 models.

5. REFERENCES

1. Asteris PG, Repapis CC, Foskolos F, Fotos A, Tsaris AK. Fundamental period of infilled RC frame structures with vertical irregularity. Structural engineering and mechanics: An international journal. 2017;61(5):663-74. 10.12989/sem.2017.61.5.663
2. Gerasimidis S, Bisbos C, Baniotopoulos C. Vertical geometric irregularity assessment of steel frames on robustness and disproportionate collapse. Journal of constructional steel research. 2012;74:76-89. 10.1016/j.jcsr.2012.02.011
3. Sarkar P, Prasad AM, Menon D. Vertical geometric irregularity in stepped building frames. Engineering Structures. 2010;32(8):2175-82. 10.1016/j.engstruct.2010.03.020
4. Di-Sarno L, Elnashai A. Seismic fragility relationships for structures. Advances in Assessment and Modeling of Earthquake Loss: Springer International Publishing Cham; 2021. p. 189-222.
5. Li S-Q. Empirical vulnerability estimation models considering updating the structural earthquake damage database. Soil Dynamics and Earthquake Engineering. 2023;169:107864. 10.1016/j.soildyn.2023.107864
6. Li S-Q, Liu H-B, Farsangi EN, Du K. Seismic fragility estimation considering field inspection of reinforced concrete girder bridges. Structure and Infrastructure Engineering. 2023:1-17. 10.1080/15732479.2023.2208565
7. Li S-Q, Liu H-B. Vulnerability prediction model of typical structures considering empirical seismic damage observation data. Bulletin of Earthquake Engineering. 2022;20(10):5161-203. 10.1007/s10518-022-01395-y
8. Schultz MT, Gouldby BP, Simm JD. Beyond the factor of safety developing fragility curves to characterize system reliability. 2010.
9. Li S-Q, Gardoni P. Empirical seismic vulnerability models for building clusters considering hybrid intensity measures. Journal of Building Engineering. 2023;68:106130. 10.1016/j.jobe.2023.106130
10. Li S-Q. Empirical resilience and vulnerability model of regional group structure considering optimized macroseismic intensity measure. Soil Dynamics and Earthquake Engineering. 2023;164:107630. 10.1016/j.soildyn.2022.107630
11. Li S-Q. Comparison of RC girder bridge and building vulnerability considering empirical seismic damage. Ain Shams Engineering Journal. 2023:102287. 10.1016/j.asej.2023.102287
12. Li S-Q, Chen Y-S. Vulnerability and economic loss evaluation model of a typical group structure considering empirical field inspection data. International Journal of Disaster Risk Reduction. 2023;88:103617. 10.1016/j.ijdrr.2023.103617
13. Ruggieri S, Uva G. Accounting for the spatial variability of seismic motion in the pushover analysis of regular and irregular RC buildings in the new Italian building code. Buildings. 2020;10(10):177. 10.3390/buildings10100177
14. Oggu P, Gopikrishna K, editors. Assessment of three-dimensional RC moment-resisting frames under repeated earthquakes. Structures; 2020: Elsevier. 10.1016/j.istruc.2020.03.039
15. Azad M, Sazzad M, Samadder N, Rahman M, editors. Effect of setback percentages in vertically irregular concrete buildings on response to earthquake. Proceedings of International Conference on Planning, Architecture and Civil Engineering; 2019.
16. Shojaei F, Behnam B. Seismic vulnerability assessment of low-rise irregular reinforced concrete structures using cumulative damage index. Advances in concrete construction. 2017;5(4):407. 10.12989/acc.2017.5.4.407
17. Ruggieri S, Porco F, Uva G, Vamvatsikos D. Two frugal options to assess class fragility and seismic safety for low-rise reinforced concrete school buildings in Southern Italy. Bulletin of Earthquake Engineering. 2021;19:1415-39. 10.1007/s10518-020-01033-5

18. Ruggieri S, Porco F, Uva G. A practical approach for estimating the floor deformability in existing RC buildings: Evaluation of the effects in the structural response and seismic fragility. Bulletin of Earthquake Engineering. 2020;18:2083-113. 10.1007/s10518-019-00774-2
19. Men JJ, Zhou Q, Shi QX. Fragility analysis method for vertically irregular reinforced concrete frame structures. Key Engineering Materials. 2008;787(400):587.
20. Kassem MM, Mohamed Nazri F, Wei LJ, Tan CG, Shahidan S, Mohd Zuki SS. Seismic fragility assessment for moment-resisting concrete frame with setback under repeated earthquakes. Asian Journal of Civil Engineering. 2019;20:465-77. 10.1007/s42107-019-00119-z
21. Nazri FM, Tan CG, Saruddin SNA. Fragility curves of regular and irregular moment-resisting concrete and steel frames. International Journal of Civil Engineering. 2018;16:917-27. 10.1007/s40999-017-0237-0
22. SAI KS, RAMA RG, MARKANDEYA RP. Seismic fragility analysis of regular and setback RCC frames-A few hypothetical case studies. 2016.
23. Ayub MA, Rizwan M, Waheed A. Damage assessment of deficient reinforced concrete setback structures. Proceedings of the Institution of Civil Engineers-Structures and Buildings. 2018;171(9):696-704. 10.1680/jstbu.17.00095
24. Mouhine M, Hilali E. Seismic vulnerability assessment of RC buildings with setback irregularity. Ain Shams Engineering Journal. 2022;13(1):101486. 10.1016/j.asej.2021.05.001
25. Mouhine M, Hilali E. Effect of setback irregularity location on the performance of RC building frames under seismic excitation. Archives of Civil Engineering. 2020;66(4). 10.24425/ace.2020.135228
26. El Janous S, El Ghoulbzouri A. Seismic Vulnerability of Irregular Reinforced Concrete Buildings Considering the Soil-structure Interaction. International Journal of Engineering. 2024;37(1):104-14.
27. Hashim A, Ali A. Structural behavior of reinforced concrete horizontally curved box beam with opening. International Journal of Engineering, Transactions A: Basics. 2022;35(4):774-83. 10.5829/IJE.2022.35.04A.17
28. Hekal GM, Kandeel K, El-Shami MM, Dawod A. Seismic fragility curves for mid-rise reinforced concrete framed structures with different lateral loads resisting systems. Concrete Research Letters. 2017;8(4). 10.20528/Cjcrl.2017.04.002
29. Sutar S, Patil R. Evaluating the Seismic Behavior of RC Structures: A Comparative Study of Framed Tube and Shear Wall Systems for Varying Numbers of Stories. 2023. 10.21203/rs.3.rs-3090596/v2
30. Computer, Structures I. Integrated finite element analysis and design of structures basic analysis reference manual. 2002.
31. Surana M, Singh Y, Lang DH. Effect of strong-column weak-beam design provision on the seismic fragility of RC frame buildings. International Journal of Advanced Structural Engineering. 2018;10:131-41. 10.1007/s40091-018-0187-z
32. Inel M, Ozmen HB. Effects of plastic hinge properties in nonlinear analysis of reinforced concrete buildings. Engineering structures. 2006;28(11):1494-502. 10.1016/j.engstruct.2006.01.017
33. Council AT. Seismic evaluation and retrofit of concrete buildings. Report No SSC 96-01: ATC-40. 1996;1.
34. Uva G, Porco F, Fiore A, Ruggieri S, editors. Effects in conventional nonlinear static analysis: Evaluation of control node position. Structures; 2018: Elsevier. 10.1016/j.istruc.2017.12.006
35. Code P. Eurocode 2: design of concrete structures-part 1–1: general rules and rules for buildings. British Standard Institution, London. 2005;668:659-68.
36. Code P. Eurocode 8: Design of structures for earthquake resistance-part 1: general rules, seismic actions and rules for buildings. Brussels: European Committee for Standardization. 2005.
37. Lang K. Seismic Vulnerability Assessment of Existing buildings, Doctoral of Technical Sciences, Department of Structural Engineering: Dissertation ETH.
38. Mazloom M, Fallah N. Seismic Vulnerability Assessment of Existing RC Moment Frames using a New Stiffness Based Damage Index. International Journal of Engineering, Transactions B: Applications. 2023;36(5):1000-11. 10.5829/IJE.2023.36.05B.16
39. Mazloom M, Fallah N. A New Approach for Seismic Damage Detection Based on Results of Pushover Analysis and Modal Based Damage Index. International Journal of Engineering, Transactions A: Basics. 2023;36(10):1892-907. 10.5829/ije.2023.36.10a.14
40. LI S, Yu T, Jia J. Empirical seismic vulnerability and damage of bottom frame seismic wall masonry structure: A case study in Dujiangyan (China) region. International Journal of Engineering, Transactions C: Aspects. 2019;32(9):1260-8. 10.5829/ije.2019.32.09c.05
41. Siqi L, Tianlai Y, Junfeng J. Investigation and analysis of empirical field seismic damage to bottom frame seismic wall masonry structure. International Journal of Engineering, Transactions B: Applications. 2019;32(8):1082-9. 10.5829/ije.2019.32.08b.04
42. Remki M, Kibboua A, Benouar D, Kehila F. Seismic fragility evaluation of existing RC frame and URM buildings in Algeria. International Journal of Civil Engineering. 2018;16:845-56. 10.1007/s40999-017-0222-7
43. Taliakula P, Prasad DV. Seismic fragility analysis of regular and vertical setback R/C frame buildings. Int Journal of Engineering Research and Applications, ISSN. 2015:2248-9622.
44. Mr H-M. Multi-hazard loss estimation methodology: Earthquake model. Department of Homeland Security, FEMA, Washington, DC. 2003:235-60.
45. Milutinovic ZV, Trendafiloski GS. Risk-UE An advanced approach to earthquake risk scenarios with applications to different european towns. Contract: EVK4-CT-2000-00014, WP4: Vulnerability of Current Buildings. 2003:1-111.
46. Barbat AH, Pujades LG, Lantada N. Seismic damage evaluation in urban areas using the capacity spectrum method: application to Barcelona. Soil Dynamics and Earthquake Engineering. 2008;28(10-11):851-65. 10.1016/j.soildyn.2007.10.006
47. Park J, Towashiraporn P, Craig JI, Goodno BJ. Seismic fragility analysis of low-rise unreinforced masonry structures. Engineering Structures. 2009;31(1):125-37. 10.1016/j.engstruct.2008.07.021

Persian Abstract

چکیده

ساختمان های بتنی مسلح نامنظم هندسی عمودی (RC) به دلیل ظاهر زیبایی شناختی و ویژگی های عملکردی آنها به طور گسترده ای در مهندسی سازه مورد استفاده قرار می گیرند. در واقع، بهبود قابلیت اطمینان و عملکرد لرزه ای آنها بسیار مهم است و حتی به یک ضرورت تبدیل شده است. این مطالعه بر اهمیت استفاده از دیوارهای برشی (SW) به عنوان یک وسیله اساسی برای تقویت این نوع ساختار تأکید می کند. بیست مدل، از جمله ده با SW و ده بدون SW، از ساختمان های متوسط با بی نظمی عقب افتاده برای این منظور در نظر گرفته شد، و تجزیه و تحلیل شکنندگی انجام شد، با استفاده از یک روش غیر خطی, برای برجسته کردن سودمندی بالقوه دیوارهای برشی برای ساختارهای نامنظم. نتایج این کار به وضوح نشان میدهد که رفتار و پاسخ دینامیکی ساختمانها با استفاده از دیوارهای برشی بهبود یافتهاست. مطالعه شکنندگی نشان می دهد که در برخی موارد احتمال آسیب کاهش می یابد، با تفاوت بیش از 13٪ برای اکثر مدل ها، و در برخی موارد تفاوت بسیار قابل توجه است, از 30٪ تا 60٪. این نشان دهنده مزایای ترکیب دیوارهای برشی در مرحله طراحی ساختمانهای نامنظم است.

IJE TRANSACTIONS B: Applications Vol. 37 No. 02, (February 2024) 439-451

International Journal of Engineering

Journal Homepage: www.ije.ir

Topical Themes and New Trends in Mining Industry: Scientometric Analysis and Research Visualization

N. V. Babyr*

Saint-Petersburg Mining University, 21 Line, 2, St. Petersburg, Russian Federation

PAPER INFO

Paper history:
Received 04 September 2023
Received in revised form 03 November 2023
Accepted 12 Novemberv2023

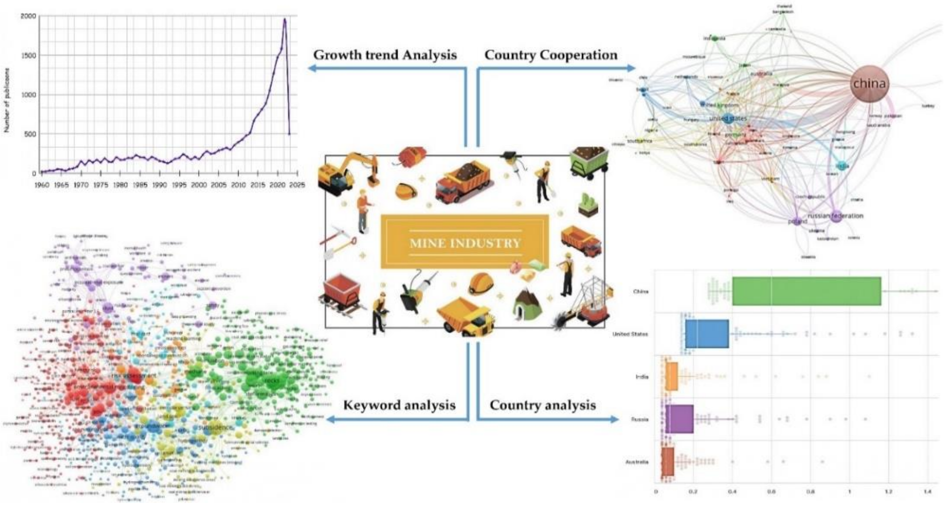
Keywords:
Scientometrics Analysis
Research Frontiers
VOSviewer
Knowledge Mapping
Mining Technologies Development

ABSTRACT

The scientometrics has become the most important tool to evaluate and analyze the performance of scientists, cooperation between universities, the impact of public funding of science on the results of national research and development, the effectiveness of education, and others. Therefore, professionals and scientists need a range of theoretical and practical tools to measure experimental data. The purpose of this article is to provide an up-to-date overview of the various tools available for scientometric analysis, including data sources, performance analysis and visualization tools. In this study performed a scientometric analysis based on 21,180 publications from the Scopus database was conducted with only published articles in mining industry between 1960 and 2023. It also revealed that 77% of the articles were published in journals and only 2.6% corresponded to review studies. Using network analysis in VOSviewer, the publications were grouped by keywords into 5 clusters with Strength. Cartographic analysis confirmed the descriptive findings and visualized the co-authorship of authors. Using bibliographic linkage analysis, the semantic relationship between authors and their associated institutions and countries was investigated (consist of 38 clusters with 404 link). Average number of citations per keyword (3.2) will allowed the most cited area is devoted to health risks. Recent studies have focused on dust and lung diseases which can pose a serious threat to the life and health of mine workers, therefore risk of coal dust explosions, which in result poses a direct risk of injury or even loss of life.

***doi**: 10.5829/ije.2024.37.02b.18*

Graphical Abstract



---

*Corresponding Author Institutional Email: babyr_nv@pers.spmi.ru (N. V. Babyr)

Please cite this article as: Babyr NV. Topical Themes and New Trends in Mining Industry: Scientometric Analysis and Research Visualization. International Journal of Engineering, Transactions B: Applications. 2024;37(02):439-51.

1. INTRODUCTION

The mining industry, a critical component of global resource extraction, finds itself at the nexus of multifaceted challenges and opportunities in the contemporary era (1-3). This sector, long renowned for its profound impact on society through the provisioning of essential raw materials, stands at a pivotal juncture influenced by intricate interplays of scientific innovation, ecological imperatives, and socio-economic dynamics (4, 5). In this context, the application of scientometric analysis and research visualization emerges as an indispensable approach to elucidate the complex, evolving landscape of mining-related research, with a profound connection to the principles of sustainable development (6, 7).

Mining, as a fundamental pillar of human civilization, has been inexorably linked to industrial progress and economic growth (8, 9). Yet, this history is punctuated by periods of environmental degradation, resource depletion, and social injustices. The modern mining industry grapples with the legacy of these issues while simultaneously seeking to chart a more sustainable and responsible path forward. Sustainable development, a cornerstone of contemporary discourse, underscores the imperative for industry sectors, including mining, to balance economic prosperity with environmental stewardship and social equity (10).

In this article, we embark on a rigorous exploration of the mining industry's scientific landscape, a journey guided by the rigorous methodologies of scientometrics and research visualization. These methodologies empower us to unravel the intricate web of scholarly publications, patents, and research outputs that constitute the mining domain. Through their lens, we gain the ability to detect patterns, quantify growth trajectories, and pinpoint pivotal moments within the rich tapestry of mining-related knowledge production (11).

Scientometrics, an empirical science in its own right, illuminates the evolutionary dynamics of mining research. It allows us to assess the proliferation of knowledge, pinpoint influential thought leaders, and identify critical nodes of intellectual convergence. By applying these systematic approaches, we set the stage for uncovering the transformative trends and thematic paradigms that coalesce around mining's journey towards sustainable development.

Scientometrics is a set of quantitative methods used to study a field of research using metadata of articles submitted to scientificometric databases (e.g., Scopus) (12). These metadata include publication title, keywords, abstract and citation records. Two main scientometric procedures are used to investigate subject areas: performance analysis and scientific mapping. Performance analysis provides a means to quantify academic output and evaluate it in terms of productivity, quality, and scholarly impact by identifying key contributors (authors, countries, organizations) and finding reliable sources of academic publications (13).

This is methods has been successfully used in the field of mining. Pouresmaieli et al. (14) analyzed 77 articles and a trend of increasing scientific interest in sustainability research in the mining industry from 1985 to 2022 was identified. Scientometric and content analysis in research Indian scientists suggested that there are few heavy metal contamination studies performed in the coal mine areas. In contrast, studies relating to assessing metal bioaccessibility and health risks in the coal mine area are relatively scarce. Finally, the gaps identified were defining the role of pH and particle size affecting bioaccessibility of metals in coal mines, the correlation between in-vitro and in vivo metal concentrations, and more clarification of Rfd values for health risk calculation (15). Likewise the result of the co-citation network study has marked the most significant authors and the highly cited sources of the database revealing Anjali and Remesan (16) as among the most cited authors with citations more than 150 in the field of our interest. The trend of publication in the research area of Water Resources showed a significant increase after 2015. The keyword occurrence map reveals that water quality studies have been extensively studied, but quantifications of the coal mining-induced changes in water regimes at river basin scales are absent (16). Also, scientometrics analysis use for trends mapping and global collaboration determination among countries, monitoring of the most productive authors, institutions, and identifying the rate of growth in recent years based on keyword analysis of authors (17).

In these publications, the authors emphasize the importance of updating research methods to reflect recent technological advances and the interdisciplinary nature of sustainable development. As technology and other areas of knowledge continue to evolve and intersect with sustainability in mining, it is important to incorporate these new developments and perspectives into future research. This will help ensure the relevance and comprehensiveness of research in the field.

A scientometric analysis provides quantitative information on the development of a given topic, including trends in results and focuses, collaborative networks, thematic research clusters, historical patterns of evolution, and collaborative citation networks (18). The advantage of scientometrics over descriptive reviews is that it can quickly identify key issues of interest and guide future research. In recent years, an increasing number of studies have applied scientometric methods to explore various aspects and obtain important information (19).

The overarching goal of this study is to utilize scientometric analysis and research visualization techniques to comprehensively explore the evolving intellectual landscape of the mining industry. The main objectives of the study:

1. To mapping and understand the current state of research, identify influential works and authors, and trace the evolution of knowledge within the mining sector.
2. To demonstrate the practical application of scientometric analysis as a robust methodology for quantifying research growth and patterns in the mining field.
3. To showcase the effectiveness of research visualization techniques, such as network graphs and co-citation maps, in visually representing the intricate tapestry of mining-related research.
4. To provide valuable insights for researchers, industry practitioners, policymakers, and stakeholders, facilitating informed decisions, future research directions, and sustainable industry practices.

2. MATERIALS AND METHOD

The scientometric data used in this study were obtained from Scopus, the largest interdisciplinary database of peer-reviewed scientific literature widely recognized and frequently used for scientometric analysis (20). Using the topical search queries «mining technologies development» OR «coal mining» OR «coal technologies development» a result of 26,565 documents from 1842 to 2023 was retrieved.

For data representativeness, the sample was limited to article and review type publications from 1960 to 2023. Another 80 publications were excluded from the sample because they were duplicated, apparently this is due to errors occurring on the side of the journal or publisher (21, 22). The largest number of duplicates came from the Geological Society Engineering Geology Special Publication, a journal that publishes engineering geology proceedings from annual conferences.

In summary, the data sample consisted of 21,180 publications, each of which was downloaded from the Scopus database and consisted of many different variables (e.g., authors, publication title, abstract, keywords, citations, references, etc.). The SCImago Journal & Country Rank (23, 24), Scopus Source list (25, 26) and Journal Citation Reports (27, 28) databases were also downloaded to analyze various indicators (Figure 1).

General information on the dataset (scientometric indicators) is summarized in the Table 1.

In addition to the traditional analysis of quantitative data on publications and citations, new, more comprehensive and informative formats for assessing scientific research are now becoming available that allow processing and visualizing very large amounts of data, such as advanced research mapping (29, 30). Such maps are typically two- or three-dimensional visualizations of the scientific landscape, consisting of topics and disciplines united by cited publications and common terminology. The degree of similarity of documents determines how far apart they will be on the map, and differences in the density of publications lead to the formation of different elements on the map, such as "mountains" and "islands" of knowledge. Mapping allows analysts to see individual scientists, organizations, grantors, and journals on a map, and to assess the extent to which an organization is engaged in research in particular areas and how that engagement has changed over time (31). In this way, the maps allow for a better understanding of the current situation, as well as the identification of key players and the most in-demand or promising new research topics (32). For creating maps based on network data and for visualizing and exploring these maps use software tool VOSviewer (13, 33).

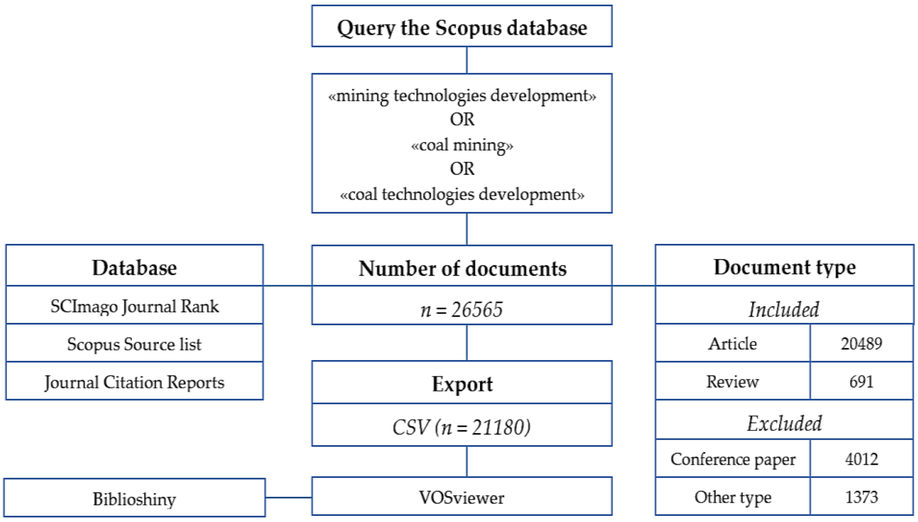


Figure 1. Methodology for determining the data set for analysis

TABLE 1. Main information

Description	Results
Timespan	1960–2023
Sources (Journals)	4170
Total documents	21180
Total citation	263865
Document average age	15.96
Article with finding	6897
Authors	46367
Single-authored documents	4660

Thus the methodology of the study can be divided into three parts:

1) Descriptive statistical analysis of the study that illustrates the current state of research in the field of mining technology development, which includes a general analysis of articles published during the sample period, as well as a network analysis of the collaboration of authors and affiliated institutions and countries. This can help researchers identify key research directions and academic leaders in the field.

2) Keyword-based coincidence analysis, which can help to understand the development of mining technology development research.

3) Detecting bursts of co-citation of references, which was conducted to analyze hot topics and emerging trends in terms of cited references.

3. RESULTS

Based on the search results, the growth of publications on this topic can be divided into 4 phases: the beginning phase from 1960 to 1985, the decline phase from 1986 to 1993, the slow growth phase from 1994 to 2012, and the fast growth phase from 2013 to 2022 (Figure 2), and Table 2 shows the most cited articles of each phase.

In the first 25 years, only 3,163 articles were published, representing only 15% of the total number of publications. From 2013 to the current moment, publications in the mining industry have more than 11,403 articles, their rapid growth can be correlated with the data on the world leaders in coal mining. According to the results of 2017, the world leader is China (share in global coal production - 44.6%), followed by India, which moved to second place (9.6%), displacing the United States (9.3%), which is now in third place; Australia (6.6%), which took fourth place; Indonesia (6.4%) - fifth place. Russia retained its sixth place in the world in coal production, its share amounting to 5.4% (5.5% in 2000) (Figure 3).

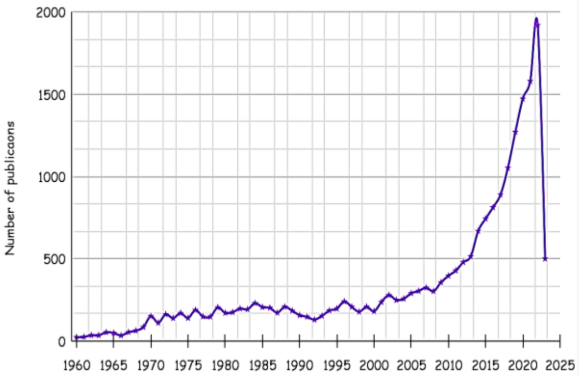


Figure 2. The annual number of publications from 1960 to 2023. The data for this article were downloaded on April 21, 2023, and therefore, only 5 months of data were included in 2023

TABLE 2. Most cited articles from 1960 to 2023

Title	Journal	h-index	Year	Cited
A thermophilic, acidophilic mycoplasme isolated from a coal refuse pile	Science	1283	1970	237
Clinically important respiratory effects of dust exposure and smoking in British coal miners	American Review of Respiratory Disease	268	1988	160
Coal mine methane: A review of capture and utilization practices with benefits to mining safety and to greenhouse gas reduction	International Journal of Coal Geology	160	2011	866
Research and development of rock mechanics in deep ground engineering	Chinese Journal of Rock Mechanics and Engineering	93	2015	469

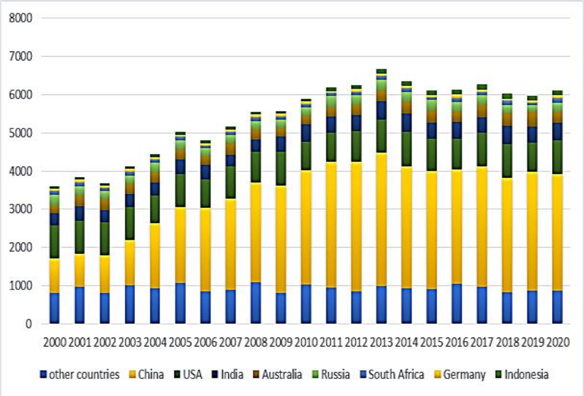


Figure 3. Coal production in major countries of the world

In the period from 2000 to 2020, the following structural changes occurred in the global coal industry under the influence of technological and price trends:
- China, India, Indonesia coal production grew at a high rate, which allowed them to occupy a market share equal to 60.7% by the end of the period, which is 22.9% more than in 2000. This growth in coal production was ensured mainly due to the increase in domestic demand, in particular in the electric and thermal power industry, as well as in machine building;
-in Australia, Russia, Kazakhstan, and South Africa, where coal production growth was driven mainly by increased coal exports, coal production increased by a combined 46.9%, but their share of the world market in 2017 declined to 16.8% from 18.8% in 2000;
- in the USA, Germany, and Poland, coal production in 2017 fell by a total of 25% compared to the 2000 level, primarily due to a reduction in the use of coal in the power sector as a result of the transition to "green" technologies. At the same time, the share of these countries in the global market more than halved and amounted to 13.3% in 2017. 13,3%.

3. 1. Analysis Journals and Journal Co-citation

A total of 4,170 scientific journals published articles between 1960 and 2023. Of these, 2,506 journals are still in existence, and 1,664 journals have ceased to exist, so for representativeness and updating of the data from the list of scientific journals, it was decided to limit the period from 2005 to 2023. Table 3 summarizes the top 20 journals in the field of mining (Only active journals were considered). In first place is the Journal of the China Coal Society with 652 papers and in second place is Coal Science and Technology with 267 papers, both published

TABLE 3. Top-20 journals publishing articles on mining industry

Source title	No. of documents	Publisher	Country	JIF	SJR	h-index	Cites per document (2 years)
Journal of the China Coal Society	652	China Coal Society	China	-	0.896	73	3,29
Coal Science and Technology	267	China Coal Society	China	-	0.55	16	1,94
Science of the Total Environment	209	Elsevier	Netherlands	9.8	1.95	317	10,94
Environmental Science and Pollution Research	193	Springer Science	Germany	5.8	0.94	154	6,18
International Journal of Coal Geology	177	Elsevier	Netherlands	5.6	1.64	160	6,03
Environmental Earth Sciences	177	Springer Verlag	Germany	2.8	0.6	141	2,88
Journal of Mining and Safety Engineering	171	China University of Mining and Technology	China	-	0.58	36	2,11
Sustainability	171	MDPI AG	Switzerland	3.9	0.66	136	4,39
Geofluids	170	Hindawi Limited	United Kingdom	1.7	0.35	64	1,67
Ugol	167	Ugol'	Russian Federation	-	0.38	15	0,78
International Journal of Environmental Research and Public Health	161	Frontiers Media S.A.	Switzerland	-	0.83	167	4,53
Energies	160	MDPI AG	Switzerland	3.2	0.63	132	3,66
Geotechnical and Geological Engineering	138	Springer Netherlands	Netherlands	1.7	0.51	71	2,02
PLoS ONE	130	Public Library of Science	United States	3.7	0.89	404	3,75
International Journal of Rock Mechanics and Mining Sciences	128	Elsevier BV	United Kingdom	7.2	1.97	187	7,2
Safety Science	120	Elsevier	Netherlands	6.1	1.43	140	7,19
Environmental Monitoring and Assessment	120	Springer Netherlands	Netherlands	3	0.63	132	3,14
Mining Informational and Analytical Bulletin	119	Publishing house Mining book	Russian Federation	-	0.42	15	1,11
International Journal of Mining Science and Technology	106	Elsevier	Netherlands	11.8	1.99	61	11,99
Journal of Liaoning Technical University	105	Liaoning Engineering Technology University Editorial	China	-	0.16	18	0,38

in China. Of the top 20 journals, 6 journals are from the Netherlands, four journals are from China, three journals are from Switzerland, Germany, United Kingdom and Russian Federation with two journals each, and one journal is from the United States. Seven journals are not indexed in the scientometric database Web of Science, and the top 10 journals have a JIF >2.0. The journal with the highest JIF and SJR is International Journal of Mining Science and Technology, and also in terms of Citations per document over two years.

3. 2. Analysis of Cooperation Analyzing the collaborative network based on country of origin can reflect the collaborative relationships between countries as well as the distribution of influential countries in the field. Figure 4 shows the scientific cooperation network obtained by VOSviewer, the node size represents the number of publications based on country/region, institution or author. According to the observation results, out of 72 countries/region, 37 published more than 25 articles, among which China ranks first with 55.96%. China is followed by Russia, USA, India, Poland, Australia, Germany, UK, Indonesia, Canada and Czech Republic with 30.3% publications in aggregate.

In addition, different countries have entered into close cooperation, which emphasizes, relevance, so the thickness and number of links connecting different nodes indicate more cooperation between developed countries and less cooperation between China and other countries. Developed countries have a larger proportion of the total number of publications, which can be explained by objective and subjective factors. As for objective factors, developed countries spend more money on scientific research. As for subjective factors, the development of mining industry as well as breakthrough technologies in developed countries is higher, which also confirms the relevance of the study.

As stated earlier, the leading position by number of publications is occupied by China - 14 073 publications, United States- 3 616 publications, Russia - 1989 publications, India- 1490 and Australia - 783 publications (Figure 5).

3. 3. Analysis of Institutional Cooperation Analyzing the collaborative network of institutions is important to understand the collaborative relationships among key institutions in mining research. Figure 6 shows that the institution collaboration network consisted of 38 clusters with 404 links. In the network, the size of the nodes represents the number of articles published by the respective institution. The thickness of the network links represents the degree of collaboration among these institutions country in mining research. Specifically, judging by the number of publications as well as the centrality index between them, China university of mining and technology has made the greatest

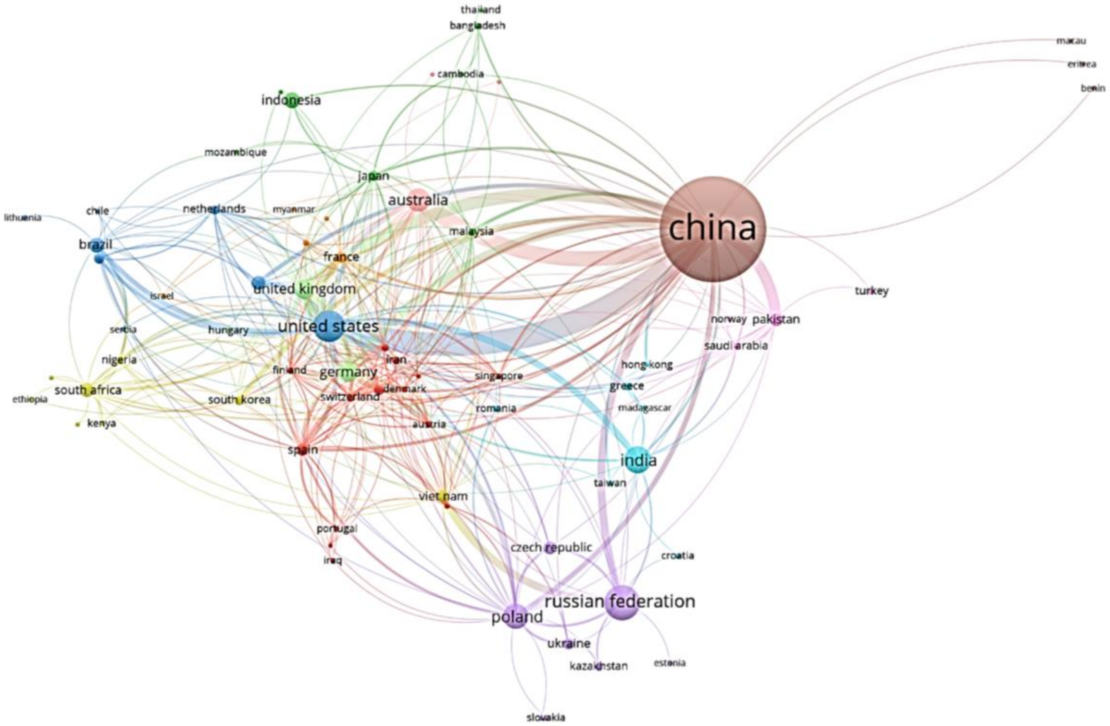


Figure 4. Country Cooperation Network on mining-related publications

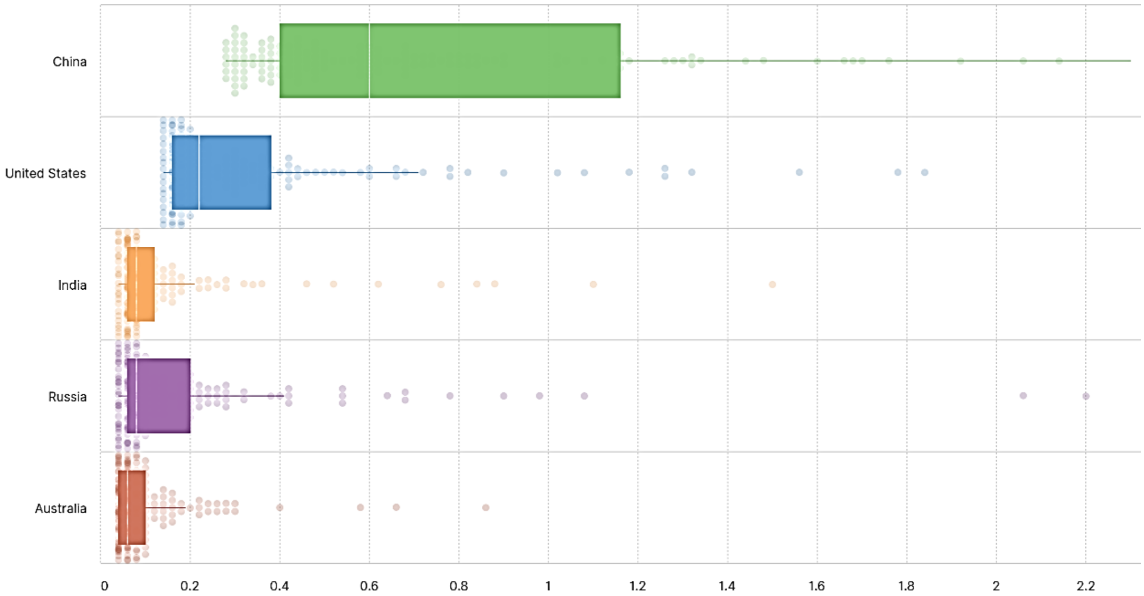


Figure 5. Top-5 countries by number of publications

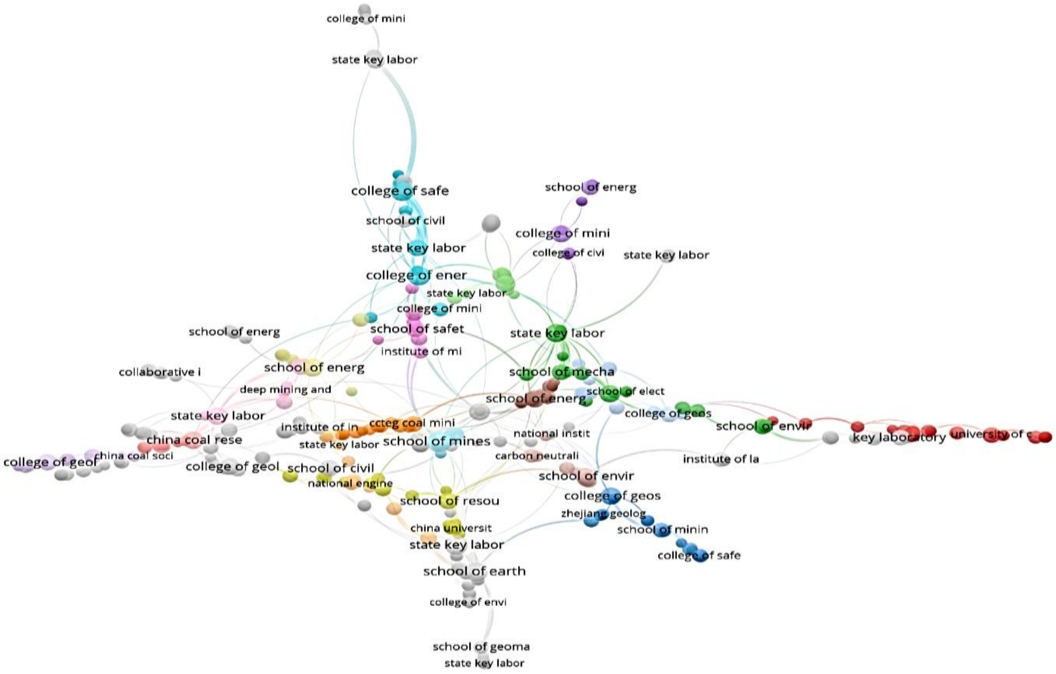


Figure 6. The institution collaboration network on mining-related publications

contribution with 44 departments or laboratories (2369 publications), followed by Shandong university of science and technology with 20 departments or laboratories (481 publications), Anhui university of science and technology with 9 departments or laboratories (375 publications), and China coal research institute with 5 departments or laboratories (228 publications).

3. 4. Analysis of the Authors Collaboration

The development and improvement of an academic scientific discipline depends on collaboration between researchers in the field. Analysis of the knowledge mapping resulting from the generation of the co-authorship network allows us to identify the most influential authors in Figure 7.

Mapping such a network can help researchers to establish collaborative relationships. According to

Scopus data, the author collaboration network created using the VOSviewer tool was disjointed and fragmented. The thickness of the links reflects the degree of collaboration between different authors. The color of the nodes reflects the time when the collaboration between authors started. The most significant collaborations in the network were those published by Wang. (34 publications) from China University of Mining and Technology (CUMT). The publication with the highest number of authors -99 (The 2022 report of the Lancet Countdown on health and climate change: health at the mercy of fossil fuels) was noteworthy. We also conducted a quantitative analysis of scientists' contributions, according to the results of which 76% of authors published only one article in their field only once, 12.5% two articles, 4.75% three articles, 2.2% four articles, and only 4.55% of scientists published 5 or more articles. Let us consider the Top 15 scientists with the most publications in the mining industry (Table 4).

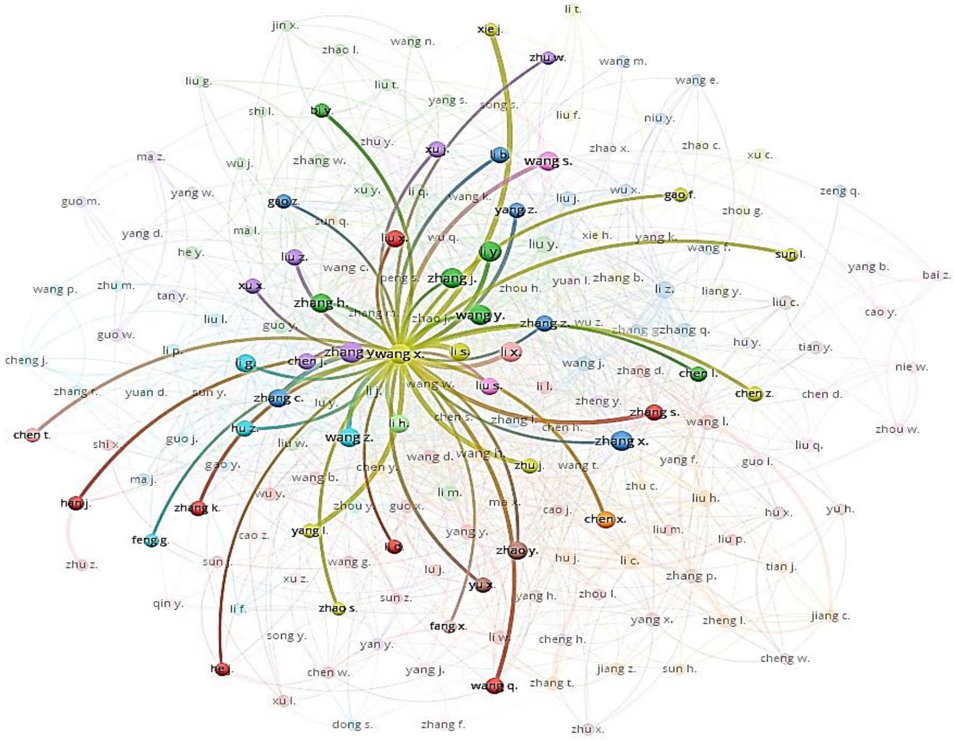


Figure 7. The author collaboration network on mining-related publications

TABLE 4. Top 15 scientists with the most publications in the mining industry

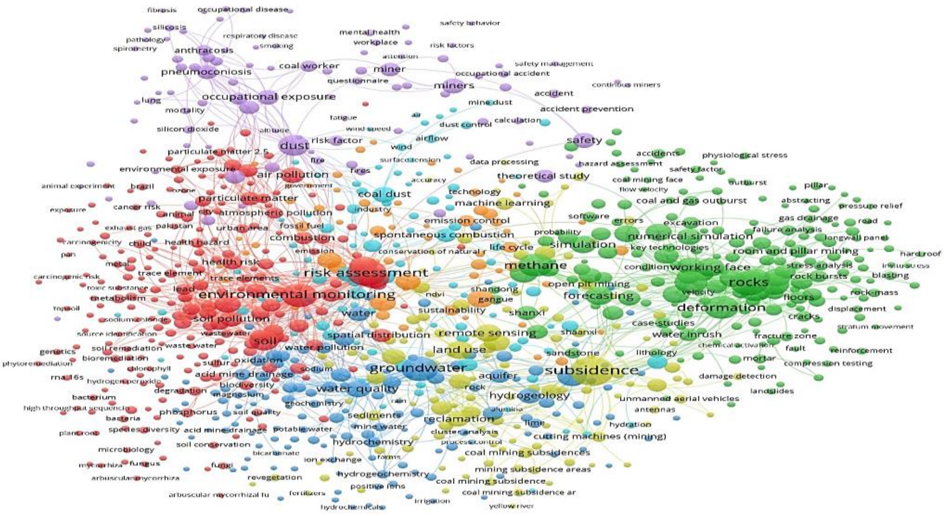
Scopus ID	Publications on the subject	h-index	Affiliation	First author	Single author
56027289900	108	44	CUMT	27	0
8257473500	83	28	CUMT	29	3
8701261700	68	19	XUST	44	0
7103314735	55	30	CUMT	10	0
12789995400	54	30	CUMT	7	0
7006661811	51	46	CUMT	16	0
7407004934	50	39	CUMT	6	2
25028121800	46	22	CUMT	8	0
36620043400	44	26	CUMT	22	0
7404915620	43	57	CUMT	10	0
7405359875	42	26	CUMT	20	0
55649649900	42	32	CUG	1	1
36112316100	42	42	CUMT	19	10
57204661209	40	29	CUMT	16	0
55588306400	40	31	CUMT	9	4

The results of the analysis indicate that institutional proximity plays an important role in scientific collaboration. In most cases, collaboration first occurred between researchers and their doctoral students, the next most common type of collaboration was collaboration between different colleagues at the same university or institute, followed by collaboration between researchers who had a past working relationship, which also confirms the lack of non-co-authored papers among scientists and their colleagues.

3. 5. Keyword Analysis

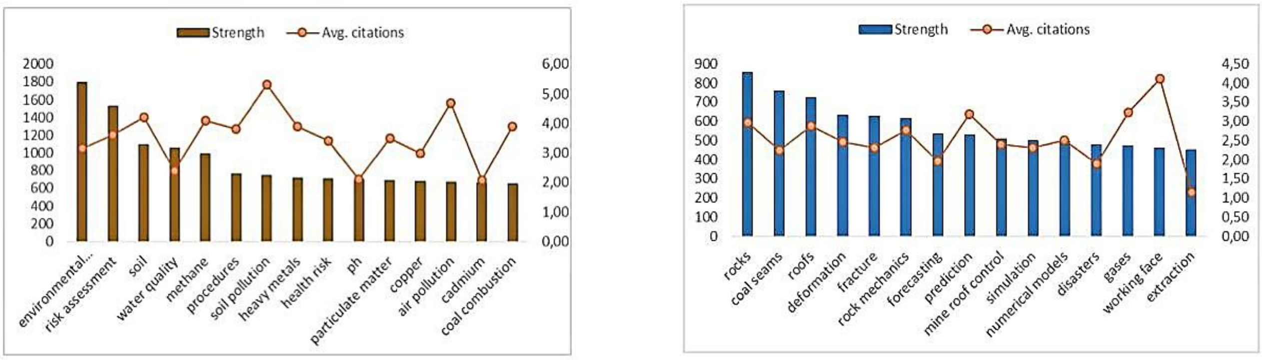
Keywords are an important indicator for understanding the main content of scientific articles. The study of emerging trends in a particular scientific field can be found by analyzing keyword evolution maps. Before proceeding, it should be noted that the documents downloaded from the Scopus database contain information on keywords only since 2005. This is due to the fact that the information contained in papers published before this date is incomplete. Thus, the keyword analysis in this study covers the period from 2005 to 2023. The keywords indicate hotspots and future trends in the research field. Keyword clustering is formed from keywords with similar research topics and each cluster is marked with frequently used keywords in articles. Keywords with a match frequency >40 were included in the keyword clustering map (Figure 8). The keywords were categorized into five clusters comprising 194 nodes. The node size and node color reflect the number of keywords and clusters. Lines of different colors indicate that two keywords occurred in the same article. Additionally, top 10 keywords were highlighted with Strength, i.e., the overall strength of the links of a given keyword with others. Cluster 1 (93 items) relates to environmental monitoring (Average citations = 3.06); cluster 2 (80 items) relates to rocks (Average citations = 1.65); cluster 3 (55 items) relates to groundwater (Average citations = 1.61); cluster 4 (42 items) relates to subsidence (Average citations = 2.66); cluster 5 (38 items) relates to health risks (Average citations = 3.2).

We summarised the top 15 keywords of each cluster to systematically understand mining industry and provide suggestions for future research (Figure 9).

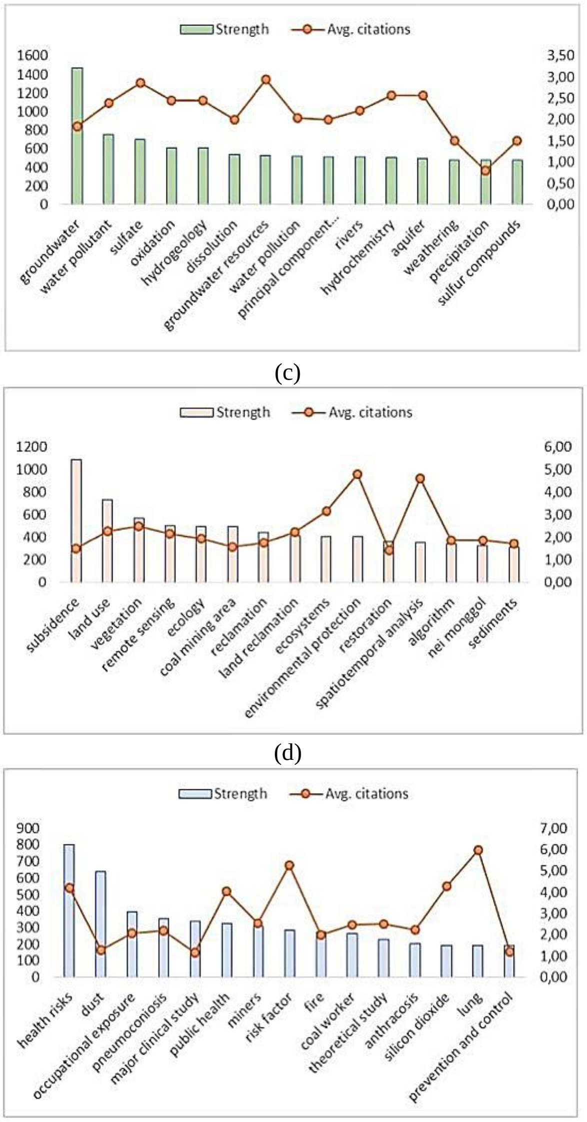


Top 10 Keywords	environmental monitoring	risk assessment	groundwater	soil	subsidence	water quality	methane	dust	aquifers	health risks
Strength	1799	1532	1464	1093	1084	1056	993	942	855	800

Figure 8. Co-occurrence of keywords and top 10 keywords with the strongest citation bursts



(a) (b)



(c)

(d)

(e)

Figure 9. The top 15 keywords of five co-occurrence keywords clusters, (a) environmental monitoring; (b) rocks; (c) groundwater; 9d) subsidence; (e) health risks

Calculating the average number of citations per keyword will allow us to determine how popular a particular area of research is in a given period. We can see that articles using the keywords risk factor (5.28), lung (6.0), working face (4.1), environmental protection (4.8), soil pollution (5.3), air pollution (4.7) gain more citations, which indicates the importance of the development of these areas.

4. DISCUSSION

A scientometric analysis and visualization of the bibliographic data were carried out and provide an insight into the most influential publication sources; most used keywords; most active and influential researchers; most active institutions, and literature with the highest impact. From the analysis, clusters were identified which grouped the literature based on the similarity of their keywords and abstract.

It was found that 54% of the articles were published between 2013 and 2023, confirming the prospectivity and prevalence of the research area. The contribution of China University of Mining and Technology to the development of this area can be separately noted, of the top 15 scientists in the world, 86% of scientists represent this organization.

The study confirmed the fact that it is impossible to assess the level of publication activity of a university by the volume of scientific research, so, according to QS World University Rankings for Engineering - Mineral and Mining, only 11 out of 25 universities are leaders in their field according to four criteria (Table 5). The first two are global surveys of academics and employers conducted by QS, which are used to assess the international reputation of institutions in each subject area. The second two indicators assess the impact of research based on the number of citations per paper and the h-index for the relevant subject (data are taken from Elsevier's Scopus database).

TABLE 5. Impact of universities on the QS World University Rankings for Engineering - Mineral and Mining

University	Country	Publucation	Rank/ Score
China University of Mining and Technology	China	3468	17/ 74.9
Pennsylvania State University	United States	164	17/ 74.9
Virginia Polytechnic Institute and State University	United States	161	28/ 70.9
University of Kentucky	United States	144	50/ 65.6
Indian Institute of Technology	India	287	25/ 71.5
Indian Institute of Technology Kharagpur	India	78	39/ 67.8
National University of Science and Technology	Russia	154	23/ 72.3
Saint Petersburg Mining University	Russia	118	3/ 83.8
The University of Queensland	Australia	179	5/ 82.1
UNSW Sydney	Australia	136	4/ 82.6
The University of Newcastle	Australia	71	31/ 69

These figures show that these institutions have had a particularly strong impact on the field of mining research. We would also like to mention the universities that carried out scientific research together with scientists from China - National university of science and technology «MISIS», Saint Petersburg state university (department of applied ecology), Saint Petersburg Mining university, Kyushu university (department of earth resources engineering), Balochistan university of information technology (department of mining engineering), Technische Universität Bergakademie Freiberg.

This dominance reflects China's position as a global leader in the mining industry, highlighting its strategic investment in research and development through funding from various finding (Table 6).

China's transformative impact on the global coal industry is undeniable. Its dominance in consumption, production, and technology innovation has ripple effects across the world, shaping the industry's trajectory. As China continues to balance its energy needs with environmental and health concerns, the global coal landscape will remain in flux, with implications for economies, energy markets, and environmental sustainability. The coal industry's future hinges on how China navigates this complex terrain and collaborates with the international community to address shared challenges.

Also the mining production process is exposed to a series of different hazards. Coal boom came at a substantial environmental and health cost. Severe air pollution, hazardous working conditions, and health-related issues prompted the leading countries in coal production to take action. Stringent environmental regulations, enforced through various policies and initiatives, have been introduced to reduce emissions, improve air quality, and address the negative impacts of coal mining and combustion. Average number of citations per keyword will allow us to determine how popular a particular area of research is in a given period. The most cited area is devoted to health risks (Average citations = 3.2). One of the main themes is the accumulation of dust and lung diseases which can pose a serious threat to the life and health of mine workers, therefore risk of coal dust explosions, which in result poses a direct risk of injury or even loss of life. The network analysis carried by keywords into confirms this fact. Recent studies have focused on both mining safety and the environment, what indicate a growing awareness of the socio-economic and health impacts of mining activities.

TABLE 6. The top documents by funding sponsor

Funding	Articles
National Natural Science Foundation of **China**	3020
Fundamental Research Funds for the Central Universities	625
National Key Research and Development Program of **China**	523
China Postdoctoral Science Foundation	299
China University of Mining and Technology	250
Natural Science Foundation of **Shandong Province**	201
Program of **Jiangsu** Higher Education Institutions	185
National Basic Research Program of **China**	156
Ministry of Education of the People's Republic of **China**	129
State Key Laboratory of Coal Resources and Safe Mining	125
Natural Science Foundation of **Jiangsu Province**	117
China Scholarship Council	104
National Science Foundation	98
European Commission	79

5. CONCLUSIONS

In the present study, a quantitative analysis of research using international scientometric databases Scopus, SJR and JCR and their data visualization tools VosViewer, Scimago Graphica and Python was conducted on the field of mining. Our results revealed the high-impact countries and institutions, journals, references, research hotspots, and key research fields in mining industry research. Since 2016 increasing publications and China leads the research on mining industry. Recent studies have focused on both mining safety and the environment. The «hot spots» of the studies include risk factor, lung, working face, environmental protection, soil pollution, air pollution. The results identified countries and institutions, journals, research hotspots, and key mining research. Thus scientometric analysis is a complete substitute for subjective literature review for academic researchers, as it allows for literature-related discoveries that would not be possible using other methods. Therefore scientometric analysis and research visualization offer valuable insights into the evolving landscape of the mining industry. From sustainable practices to technological advancements, these methods help us understand the trends and themes shaping the future of mining. As the industry continues to adapt to global challenges, scientometrics will remain a crucial tool for monitoring and guiding its development.

5. REFERENCES

1. Tsyglianu P. Engineering projects in the russian fuel and energy complex: actual problems, factors and recommendations for development, Ugol', 3. doi. 2023;10:0041-5790. 10.18796/0041-5790-2023-3-45-51

2. Zhdaneev OV. Technological sovereignty of the Russian Federation fuel and energy complex. Записки Горного института. 2022;258:1049-66. 10.31897/PMI.2022.107
3. Nikolaichuk L, Ignatiev K, Filatova I, Shabalova A. Diversification of Portfolio of International Oil and Gas Assets using Cluster Analysis. International Journal of Engineering, Tansaction A: Basics. 2023;36(10):1783-92. 10.5829/IJE.2023.36.10A.06
4. 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
5. Kuili S, Sastry VR. A Numerical Modelling Approach to Assess Deformations of Horseshoe Cavern on Account of Rock Mass Characteristics and Discontinuities. International Journal of Engineering, Tansaction A: Basics. 2023;36(7):1259-68. 10.5829/ije.2023.36.07a.07
6. Yurak VV, Dushin AV, Mochalova LA. Vs sustainable development: Scenarios for the future. Записки Горного института. 2020;242:242-7. 10.31897/PMI.2020.2.242
7. Litvinenko V, Tsvetkov P, Molodtsov K. The social and market mechanism of sustainable development of public companies in the mineral resource sector. Eurasian Min. 2020;2020:36-41. 10.17580/em.2020.01.07
8. Kazanin O, Sidorenko A, Drebenstedt C. Intensive underground mining technologies: Challenges and prospects for the coal mines in Russia. Acta Montanistica Slovaca. 2021;26(1). 10.46544/AMS.v26i1.05
9. Amirshenava S, Osanloo M, Esfahanipour A. Development of open-pit mine reclamation cost estimation models: A regression-based approach. International Journal of Engineering, Transactions B: Applications. 2021;34(11):2467-75. 10.5829/ije.2021.34.11b.10
10. Minin I, Nedyalkov P, Hristova T, Kraychev E, Dimov E, editors. Conceptual options for solving the problem of icing rollers and drums of belt conveyors. IOP Conference Series: Earth and Environmental Science; 2023: IOP Publishing. 10.1088/1755-1315/1156/1/012001
11. Cherepovitsyn AE, Tsvetkov PS, Evseeva OO. Critical analysis of methodological approaches to assessing sustainability of arctic oil and gas projects. Записки Горного института. 2021;249:463-78. 10.31897/PMI.2021.3.15
12. Mingers J, Leydesdorff L. A review of theory and practice in scientometrics. European journal of operational research. 2015;246(1):1-19. 10.1016/j.ejor.2015.04.002
13. Moral Muñoz JA, Herrera Viedma E. Software tools for conducting bibliometric analysis in science: An upto-date review. 2020. 10.3145/epi.2020.ene.03
14. Pouresmaieli M, Ataei M, Qarahasanlou AN. A scientometrics view on sustainable development in surface mining: Everything from the beginning. Resources Policy. 2023;82:103410. 10.1016/j.resourpol.2023.103410
15. Kumari M, Bhattacharya T. A review on bioaccessibility and the associated health risks due to heavy metal pollution in coal mines: Content and trend analysis. Environmental Development. 2023:100859. 10.1016/j.envdev.2023.100859
16. Anjali K, Remesan R. Exploring the last 50 years of Indian research on the impact of coal mining using bibliometric analysis with an overview of water-related impacts. Environmental Science and Pollution Research. 2023;30(6):16449-63. 10.1007/s11356-022-23381-2
17. Mostafa Hatami A, Sabour MR, Joshaghani A. Research trends on ash stabilization in the pavement during 2002–2021. Environmental Science and Pollution Research. 2023;30(1):1611-21. 10.1007/s11356-022-22250-2
18. Li J, Goerlandt F, Reniers G. An overview of scientometric mapping for the safety science community: Methods, tools, and framework. Safety Science. 2021;134:105093. 10.1016/j.ssci.2020.105093
19. Akinlolu M, Haupt TC, Edwards DJ, Simpeh F. A bibliometric review of the status and emerging research trends in construction safety management technologies. International Journal of Construction Management. 2022;22(14):2699-711. 10.1080/15623599.2020.1819584
20. Udara Willhelm Abeydeera LH, Wadu Mesthrige J, Samarasinghalage TI. Global research on carbon emissions: A scientometric review. Sustainability. 2019;11(14):3972. 10.3390/su11143972
21. Franceschini F, Maisano D, Mastrogiacomo L. The museum of errors/horrors in Scopus. Journal of Informetrics. 2016;10(1):174-82. 10.1016/j.joi.2015.11.006
22. Buchanan RA. Accuracy of cited references: The role of citation databases. College & Research Libraries. 2006;67(4):292-303. 10.5860/crl.67.4.292
23. Guerrero-Bote VP, Moya-Anegón F. A further step forward in measuring journals’ scientific prestige: The SJR2 indicator. Journal of informetrics. 2012;6(4):674-88. 10.1016/j.joi.2012.07.001
24. Roldan-Valadez E, Salazar-Ruiz SY, Ibarra-Contreras R, Rios C. Current concepts on bibliometrics: a brief review about impact factor, Eigenfactor score, CiteScore, SCImago Journal Rank, Source-Normalised Impact per Paper, H-index, and alternative metrics. Irish Journal of Medical Science (1971-). 2019;188:939-51. 10.1007/s11845-018-1936-5
25. Mishra M, Singh R. Regional science, Regional Planning, and the Global south. Practices in Regional Science and Sustainable Regional Development: Experiences from the Global South. 2021:3-21. 10.1007/978-981-16-2221-2_1
26. Vit Machacek, editor Globalization of Science: Evidence from Authors in Academic Journals by Country of Origin. 17th International Conference on Scientometrics and Informetrics; 2020. 10.48550/arXiv.2112.02672
27. Montero-Díaz J, Cobo M-J, Gutiérrez-Salcedo M, Segado-Boj F, Herrera-Viedma E. A science mapping analysis of ‘Communication’WoS subject category (1980-2013). Comunicar: Revista Científica de Comunicacíon y Educacíon. 2018;26(55):81-91. 10.3916/C55-2018-08
28. Guan J, Yan Y, Zhang JJ. The impact of collaboration and knowledge networks on citations. Journal of Informetrics. 2017;11(2):407-22. 10.1016/j.joi.2017.02.007
29. Cobo MJ, López-Herrera AG, Herrera-Viedma E, Herrera F. Science mapping software tools: Review, analysis, and cooperative study among tools. Journal of the American Society for information Science and Technology. 2011;62(7):1382-402. 10.1002/asi.21525
30. Ravikumar S, Agrahari A, Singh SN. Mapping the intellectual structure of scientometrics: A co-word analysis of the journal Scientometrics (2005–2010). Scientometrics. 2015;102:929-55. 10.1007/s11192-014-1402-8
31. Hou Q, Mao G, Zhao L, Du H, Zuo J. Mapping the scientific research on life cycle assessment: a bibliometric analysis. The International Journal of Life Cycle Assessment. 2015;20:541-55. 10.1007/s11367-015-0846-2
32. Katsurai M, Ono S. TrendNets: mapping emerging research trends from dynamic co-word networks via sparse representation. Scientometrics. 2019;121(3):1583-98. 10.1007/s11192-019-03241-6
33. Ding Y, Rousseau R, Wolfram D. Visualizing bibliometric networks. Measuring Scholarly Impact: Methods and Practice. 2014:285-320. 10.1007/978-3-319-10377-8_13

Persian Abstract

چکیده

علم سنجی به مهم ترین ابزار برای ارزیابی و تحلیل عملکرد دانشمندان، همکاری بین دانشگاه ها، تأثیر بودجه عمومی علم بر نتایج تحقیق و توسعه ملی، اثربخشی آموزش و غیره تبدیل شده است. بنابراین، متخصصان و دانشمندان برای اندازه‌گیری داده‌های تجربی به طیف وسیعی از ابزارهای نظری و عملی نیاز دارند. هدف از این مقاله ارائه یک نمای کلی به روز از ابزارهای مختلف موجود برای تجزیه و تحلیل علم سنجی، از جمله منابع داده، تجزیه و تحلیل عملکرد و ابزارهای تجسم است. در این مطالعه تجزیه و تحلیل علم سنجی بر اساس 21180 مقاله از پایگاه داده اسکوپوس انجام شد که تنها با مقالات منتشر شده در صنعت معدن بین سال های 1960 تا 2023 انجام شد. همچنین نشان داد که 77 درصد از مقالات در مجلات منتشر شده و تنها 2.6 درصد با مطالعات مروری مطابقت دارد. . با استفاده از تحلیل شبکه در VOSviewer، نشریات بر اساس کلمات کلیدی در 5 خوشه با Strength گروه بندی شدند. تجزیه و تحلیل نقشه‌برداری، یافته‌های توصیفی را تأیید کرد و هم‌نویسی نویسندگان را به تصویر کشید. با استفاده از تحلیل پیوند کتابشناختی، رابطه معنایی بین نویسندگان و موسسات و کشورهای مرتبط با آنها بررسی شد (شامل 38 خوشه با 404 پیوند). میانگین تعداد نقل‌قول‌ها به ازای هر کلمه کلیدی (3.2) اجازه می‌دهد بیشترین استناد منطقه به خطرات سلامتی اختصاص داده شود. مطالعات اخیر بر گرد و غبار و بیماری های ریوی متمرکز شده است که می تواند تهدیدی جدی برای زندگی و سلامت کارگران معدن باشد، بنابراین خطر انفجار گرد و غبار زغال سنگ، که در نتیجه خطر مستقیم آسیب یا حتی از دست دادن جان افراد را به همراه 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.

International **J**ournal of **E**ngineering, P.O. Box 13145-1659, Tehran, I.R. Iran
Materials and Energy Research Center, 5 Ahuramazda St., Alvand Ave., Argentina Sq., Tehran, Islamic Republic of Iran
www.merc.ac.ir - www.emerc.ac.ir

Effect of Current Density on Magnetic and Hardness Properties of Ni-Cu Alloy Coated on Al via Electrodeposition

by Cahaya Rosyidan FTKE

IJE TRANSACTIONS B: Applications Vol. 37 No. 02, (February 2024) 213-223

International Journal of Engineering

Journal Homepage: www.ije.ir

Effect of Current Density on Magnetic and Hardness Properties of Ni-Cu Alloy Coated on Al via Electrodeposition

C. Rosyidan[a,b], B. Kurniawan*[a], B. Soegijono[c], V. G. Vidia Putra[d], D. R. Munazat[a], F. B. Susetyo[e]

[a] *Department of Physics, Universitas Indonesia, Depok, Indonesia*
[b] *Department of Petroleum, Universitas Trisakti, Jakarta, Indonesia*
[c] *Department of Geoscience, Universitas Indonesia, Depok, Indonesia*
[d] *Plasma and Nanomaterial Research Group, Politeknik STTT Bandung, Bandung, Indonesia*
[e] *Department of Mechanical Engineering, Universitas Negeri Jakarta, Jakarta, Indonesia*

PAPER INFO

Paper history:
Received 28 July 2023
Received in revised form 08 Septemebr 2023
Accepted 30 September 2023

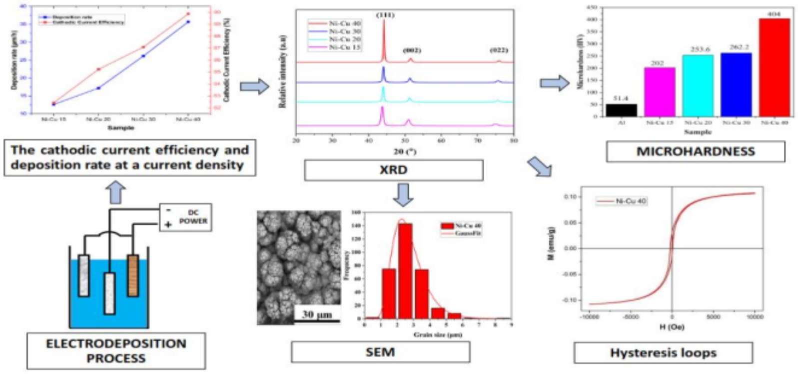
Keywords:
Cathodic Current Efficiency
Microhardness
Ni-Cu Coating
Vibrating Sample Magnetometer

ABSTRACT

Nickel (Ni)-rich single-phase nickel-copper (Ni-Cu) alloy coatings were produced on aluminum (Al) substrates by electrodeposition in stabilized citrate baths. Electrodeposition experiments were performed at four different current densities. Increasing the current density resulted in the metal deposition rate increasing faster than the hydrogen evolution rate; thus, the cathodic current efficiency increased. The crystal systems of the Ni-Cu alloys were face center cubic (fcc), with the (111) plane as the preferred crystal plane. Scanning electron microscopy with energy dispersive X-ray spectroscopy (SEM-EDS) measurements showed that the Ni content in the coating increased with increasing current density. The Ni-Cu 40 sample had the most Ni content and showed a homogeneous and compact morphology. It was found that the higher the concentration of Ni in the solution, the smaller the grain size. Measurements recorded with a vibrating sample magnetometer (VSM) showed that the Ni-Cu 40 sample provided magnetic saturation, with the highest value being 0.108 emu/g. The microhardness method produced 404 HV on the Ni-Cu 40 sample. In conclusion, higher current densities were associated with a higher Ni composition and increased thickness, which were responsible for the increases in the magnetic properties and hardness.

doi: *10.5829/ije.2024.37.02b.01*

Graphical Abstract



---

*Corresponding author email: *budhy.kurniawan@sci.ui.ac.id*. (B. Kurniawan)

Please cite this article as: Rosyidan C, Kurniawan B, Soegijono B, Vidia Putra VG, Munazat DR, Susetyo FB. Effect of Current Density on Magnetic and Hardness Properties of Ni-Cu Alloy Coated on Al via Electrodeposition. International Journal of Engineering, Transactions B: Applications. 2024;37(02):213-23.

NOMENCLATURE

C_e	Cathodic current efficiency	W_i	Initial weight of the substrate
σ	Lattice strain	W_f	Weight
W_s	Final weight of the substrate	I	Total current
W_m	Ratio of the final weight of the substrate	t	Deposition time
μ	Texture coefficient of the unique plane	F	Faraday's constant
$I(hkl)$	Measured intensity	f_{ni}	Nickel deposit weight ratio
m_{cu}	Copper's atomic weight	m_{ni}	Nickel's atomic weight

1. INTRODUCTION

Researchers widely study nickel (Ni) and copper (Cu) alloys as engineering materials due to their unique mechanical, magnetic, and anti-corrosion properties (1, 2). Ni-Cu alloys are known as monel in the industry and are typically comprised of 70 wt% Ni and 30 wt% Cu (3). These alloys have outstanding capabilities in acidic and alkaline environments (4). Ni-Cu alloys are single-phase alloys throughout their composition on the phase diagram, and these alloys formed because Ni and Cu are fully soluble in their solid and liquid states (5). Ni and Cu both have a face center cubic (fcc) crystal structure, and they have almost similar electronegativity and atomic radii (6, 7).

Given that conventional casting as a manufacturing method for monel results in substantial production costs and that Ni-Cu alloy coated on aluminum (Al) has potential as a replacement for monel as a bulk material (8), Al-based metals have received considerable attention (9, 10). They are lightweight and demonstrate high resistance to wear and corrosion and a high strength–stiffness combination (11).

Several techniques have been proposed to successfully modify the surface morphology and chemical composition, including sol-gel, chemical etching, chemical vapor deposition, thermal embossing, and electrodeposition (12, 13). The electrodeposition technique is a cost-effective, scalable, and easy-to-control process for coating Ni-Cu alloy (14). Specific methods have also been developed to determine the structure, morphology, and phase composition of the coated Ni-Cu alloys (15, 16). Goranova et al. (17) investigated how changing the concentration of Ni ions and the current density affected the structure and composition of Ni-Cu alloys formed by electrodeposition in alkaline citrate baths. Higher concentrations of Ni ions in the bath led to notably smoother deposits and enhanced current efficiency. However, producing a uniform Ni-Cu coating can be challenging due to the difference in reduction potential between Ni and Cu. The reduction potential of Ni atoms is -0.25 V vs. SHE, and that of Cu atoms is +0.34 V vs. SHE (18). As a result, controlling the concentrations of Ni and Cu is vital. Complexing agents must be added to narrow the potential difference between Ni and Cu. The most frequently used complexing agent is citrate due to its low toxicity, cost-effectiveness, and buffering characteristics (19).

The electrodeposition process affects the physical properties of the resultant Ni-Cu alloy, as does the current density. A high current density causes the crystal plane to be oriented in the (111) plane, the lattice size to be smaller, and the atomic distance to be less (20). The grain size becomes smaller when the current density is high, and the morphological shape resembles that of a cauliflower (21). As a result of a high current density, the coating will be thicker, and the composition of the Ni weight fraction will also be higher. The amount of Ni deposited on the substrate and the thickness of the coating both have an impact on the product's magnetic properties (22). In addition, a smaller grain size results in an increase in hardness (23). Karunakaran et al. (24) reported a hardness of 153 HV when the current density was 40 mA/cm^2, and Karunakaran and Pugazh (25) Vadivu reported a magnetic saturation value of 0.0004 emu/g at 40 mA/cm^2. Nevertheless, the researchers did not examine the impact of the coating electrodeposition factors, structure, and morphology on the magnetic and hardness properties.

The aims of this research were 1) to produce a Ni-rich Ni-Cu alloy coating on Al via electrodeposition and 2) to investigate the link between magnetic and hardness properties and the coating's microstructure and surface morphology. We varied the current density, and the process was conducted at room temperature. We examined the influence of various process variables on the cathodic current efficiency, structure, morphology, composition, grain size, and thickness of the produced coatings. Finally, the magnetic properties and hardness of the coatings were investigated.

2. MATERIAL AND METHODS

2. 1. Material and Electrodeposition Process

The chemical composition of the Al substrate (cathode) used was Fe = 1.63 wt%, Mg = 1.49 wt%, and Al = 96.88 wt%. The chemical composition of the Ni (anode) used was Al = 0.02 wt%, Ca = 0.04 wt%, Fe = 0.23 wt%, Y = 1.61 wt%, Zr = 0.04 wt%, Nb = 0.05 wt%, and Ni = 98.01 wt%. The chemical composition of the Cu (anode) used was P = 0.22 wt%, Cd = 0.684 wt%, Si = 0.137 wt%, and Cu = 98.959 wt%. The Al was cleaned from the oxide coating with sandpaper before deposition using DELTA D68H for 5 min. Ni-Cu electrodeposition was carried out using a SANFIX 305 E DC power supply. The samples produced using a current density of

15 mA/cm^2, 20 mA/cm^2, 30 mA/cm^2, and 40 mA/cm^2 were designated as Ni-Cu 15, Ni-Cu 20, Ni-Cu 30, and Ni-Cu 40, respectively. Table 1 summarized the bath composition and deposition parameters.

2. 2. Characterization The deposition rate was calculated using the previously reported method (26). The following formula, Equation 1, was used to calculate the efficiency of the cathodic current (27):

$$C_e = {}^{W_m}\!/_{W_f} \qquad (1)$$

W$_m$ and W$_f$ were calculated using Faraday's law, as shown in Equations 2 and 3.

$$W_m = W_s - W_i, \qquad (2)$$

$$W_f = \{(m_{ni}/2) * f_{ni} + (m_{cu}/2) * f_{cu}\} * I.\frac{t}{F}. \qquad (3)$$

The crystal structure of the Ni-Cu coating was determined using X-ray diffraction (XRD-PANalytical Aeris Instrument Suit) (Cu-Kα radiation, λ = 0.15418 nm). XRD data were collected from 20° to 80° with a step size of 0.020°. The Materials Analysis Using Diffraction (MAUD) program was used to determine the crystal parameters of the sample after Rietveld refinement. The preferential crystallite orientation was determined from the texture coefficient μ, as shown in Equation 4 (28):

$$\mu = \frac{I(hkl)/I_0(hkl)}{(\frac{1}{N})\sum[I(hkl)/I_0(hkl)]}, \qquad (4)$$

Based on XRD results, the lattice strain σ was calculated using Equation (5) (29):

$$\sigma = \left(\frac{\beta}{4 \times \tan\theta}\right) \qquad (5)$$

SEM-EDS (Thermofisher Quanta 650 EDAX EDS Analyzer) with 1000× magnification was used to analyze the surface morphology of the Ni-Cu coatings. EDS was used to determine the chemical composition of the coatings, and the statistical distribution of grain sizes was calculated using ImageJ software. The cross sections of the coated samples were also examined to assess how the current density and CCE affected the coating thickness. Measurement of magnetic properties was conducted using a vibrating sample magnetometer (VSM, Oxford 1.2H). The hardness of the Ni-Cu coatings was measured using a MicroMct® 5100 Series Microindentation Hardness Tester. The ATM E384 standard was used for the tests, which were performed with a load of 100 g for 10 s at five places.

TABLE 1. Bath composition and deposition parameters

Bath composition and condition	Quantity
NiSO$_4$.6H$_2$O (Merck)	0.5 M
CuSO$_4$.5H$_2$O (Merck)	0.04 M
Na$_3$C$_6$H$_5$O$_7$ (Merck)	0.2 M
pH	4.2
Temperature	25 °C
Deposition time	1 h

3. RESULTS AND DISCUSSION

3. 1. Cathodic Current Efficiency and Deposition Rate Figure 1 depicts the relationship between the average CCE and deposition rate in the citrate electrolyte bath. The CCE was found to be high, with a value of 82–89%. The basic concept of current efficiency can be understood as the fraction of total current used for metal plating (30). Apart from metal deposition, hydrogen evolution is the only other necessary process that must occur on the substrate surface. Ni and Cu precipitation are both antagonistic to the hydrogen evolution reaction. In this study, the metal deposition rate increased faster than the hydrogen evolution rate when the current density increased from 15 to 40 mA/cm^2; hence, the CCE increased. The highest CCE (89.96%) was associated with the Ni-Cu 40 sample, and the lowest CCE (82.55%) was associated with the Ni-Cu 15 sample. Basori et al. and Syamsuir et al. (31) found that the deposition rate and current efficiency are considered linear.

3. 2. Structural Properties Figure 2(a) depicts the XRD spectra of the Ni-Cu alloy samples produced at various current densities. According to the XRD analysis results, each Ni-Cu alloy sample consisted of a single phase with an fcc structure. The evolution of the lattice parameters of the Ni-rich (111) phase over the range of current densities is shown in Figure 2(b). The

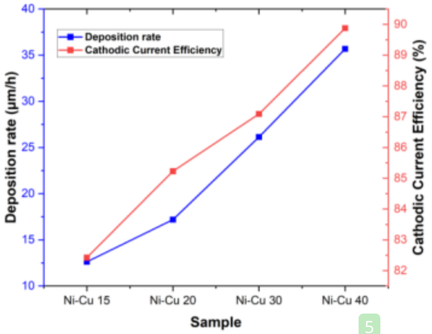
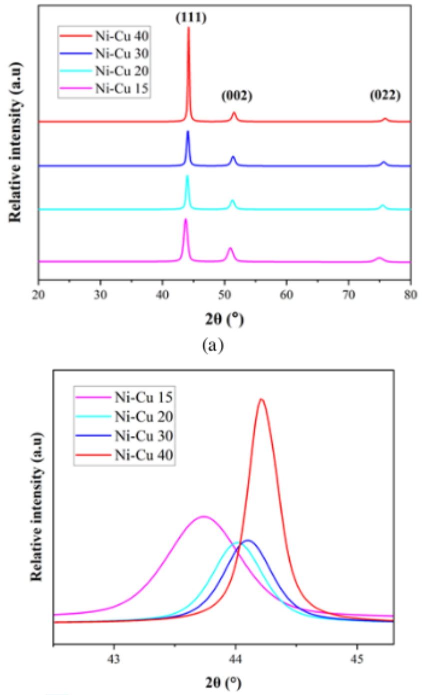


Figure 1. The cathodic current efficiency and deposition rate at a current density of 15, 20, 30, and 40 mA/cm^2



(a)

(b)

Figure 2. (a) X-ray diffraction spectra of Ni-Cu coatings electrodeposited at various current densities and (b) the extended view of the Ni-Cu (111) plane, showing peak shifts

peaks of each sample were found between the peaks of the fcc of 2θ = 43.3° for pure Cu and 2θ = 44.5° for pure Ni (32, 33). As the Ni content of the alloy coating increased, the diffraction angle also increased.

Crystal size calculation using MAUD resolved refinement was used to determine the size of the crystallites in the Ni-Cu alloys, and the results (Table 2) show that the crystallite size of the Ni-Cu coating ranged from approximately 24 to 50 nm. The crystallite size of a pure Ni layer is 60 nm, meaning that the Ni-Cu alloys had smaller crystallite sizes than a pure Ni layer. This result is also similar to that obtained by Li et al. (34). In contrast to the typical watt-Ni coating, we found that the presence of a sodium citrate complexing agent resulted in a finer crystallite size. This is consistent with the findings of Sarac and Baykul (35), who observed that Cu atoms affect grain refinement in Ni-Cu alloys. Cu atoms can restrain the surface diffusion of Ni atoms during the deposition process and inhibit the growth of crystallites.

The evolution of the crystallographic orientation of the Ni-Cu coatings produced with varying current densities is shown in detail in Figure 3. It can be observed that a strong (111) fiber texture appeared in all the samples, while the (002) texture gradually decreased as the current density increased.

The μ values of different crystal planes are also used to evaluate the degree of crystallographic orientation (34). Moreover, the Ni-Cu coating electrodeposited at the current density of 40 mA/cm^2 was found to have a strong (111) texture.

The texture coefficient for every preference was calculated using Equation 4 to ascertain the preferred crystal orientation direction of each Ni-Cu alloy obtained at the various current densities, and the results are shown in Table 3 (32).

It seems that the texture coefficient was also dependent on the peak current, and the preferred orientation was the (111) plane. Li et al. (34) found that the higher the current density, the more dominant the (111) plane. The findings suggest that the (111) crystallographic orientation was preferable for all the Ni-Cu coatings electrodeposited at the tested current densities.

The lattice strain of the prepared coatings was determined using Equation 5, and Figure 4 illustrates the changes in the crystal size and lattice strain of the Ni-Cu coatings based on the current density applied in the plating bath.

The crystal size increased and the lattice strain decreased as the current density increased (36). This result aligns with that obtained by Devi et al. (33), who showed that the higher the current density, the more the

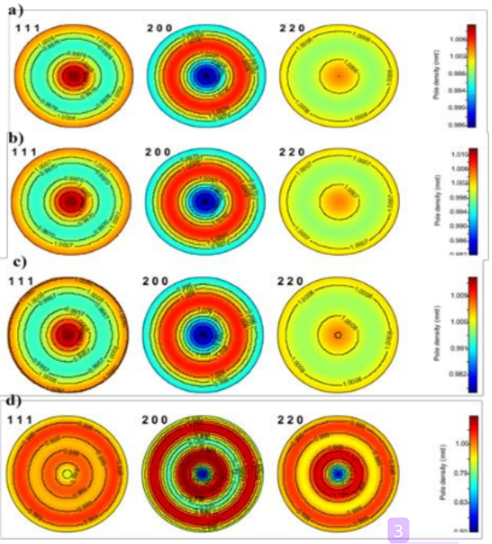


Figure 3. The simulated 2D pole figures for the (a) Ni-Cu 15, (b) Ni-Cu 20, (c) Ni-Cu 30, and (d) Ni-Cu 40 samples

TABLE 2. Parameters of the Ni-Cu alloys after Rietveld refinement using MAUD

	Sample			
Parameter	**Ni-Cu 15**	**Ni-Cu 20**	**Ni-Cu 30**	**Ni-Cu 40**
Crystal structure	Cubic fcc			
Space group	Fm-3m			
Lattice constant (Å) a = b = c	3.582	3.560	3.554	3.545
Volume (Å^3)	45.975	45.152	44.905	44.557
d-spacing (Å)	1.791	1.780	1.695	1.691
Crystallite size (nm)	24.55	26.82	32.29	50.78
Rwp (100%)	3.640	4.484	5.394	5.139
GOF	1.94	1.72	2.06	2.03
Lattice strain	0.663	0.453	0.431	0.276

TABLE 3. Texture coefficient analysis of Ni-Cu deposits

Sample	μ (hkl)		
	[111]	[002]	[022]
Ni-Cu 15	1.12	0.86	0.77
Ni-Cu 20	1.17	0.74	0.81
Ni-Cu 30	1.18	0.74	0.78
Ni-Cu 40	1.48	0.34	0.32

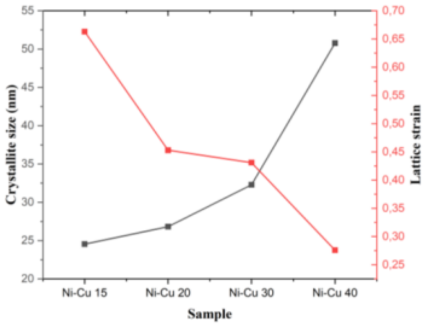


Figure 4. The lattice strain and crystallize size of the Ni-Cu alloys coated on Al at various current densities

crystal size increased. A possible reason for this is that the composition of Ni increases as the current density increases.

3. 3. Surface Morphology EDS was used to determine the elemental composition of the Ni-Cu coatings, and the results are shown in Figure 5 and Table 4. Cu and Ni were the only elements present in the deposits. The alloy composition was influenced by the current density: as the current density increased, the Cu content decreased. This phenomenon can be caused by the [Ni^{2+}]/[Cu^{2+}] ratio in the bath, which changes the composition of Ni and Cu. Goranova et al. discovered that as the Cu content of deposits decreased, so did the CCE (37). This phenomenon occurs because of the orderly deposition of Ni and Cu (38). In regular deposition, increasing the current density leads to an increased proportion of less noble metals in the deposited material (39). In our scenario, Ni is the less noble metal. As a result, when the current density was higher, the deposits were richer in Ni.

Another notable feature was that as the current density increased, the peak shifted to the right (i.e., to higher 2θ values). The change in the alloy composition may have also caused this peak shift. The Ni concentration increased as the current density increased (see Table 4). Because Ni and Cu combine to produce a single-phase alloy, the diffraction peak shifted toward that of pure Ni as the Ni percentage increased. This finding is similar to the observations of Goranova et al. (17), who found that the fcc reflection for Ni-rich Ni-Cu alloy deposits shifted to the right as the Ni concentration increased. Indeed, it is logical to expect the Ni-Cu alloy peak to shift as the Ni content increase.

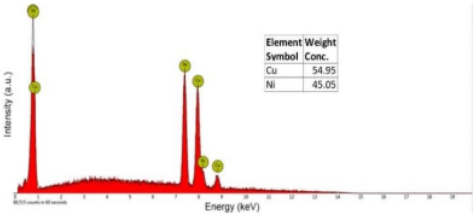


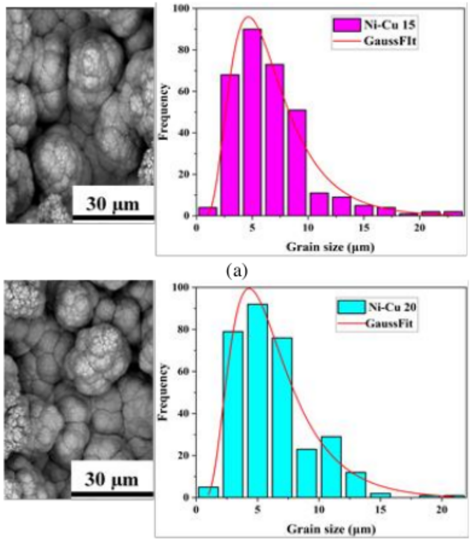
Figure 5. Eds graph of the Ni-Cu 15

TABLE 4. The chemical composition of Ni-Cu coatings on the Al substrate prepared at different current densities

Sample	Cu, wt%	Ni, wt%
Ni-Cu 15	54.95	45.05
Ni-Cu 20	39.83	60.17
Ni-Cu 30	29.80	70.20
Ni-Cu 40	19.64	80.36

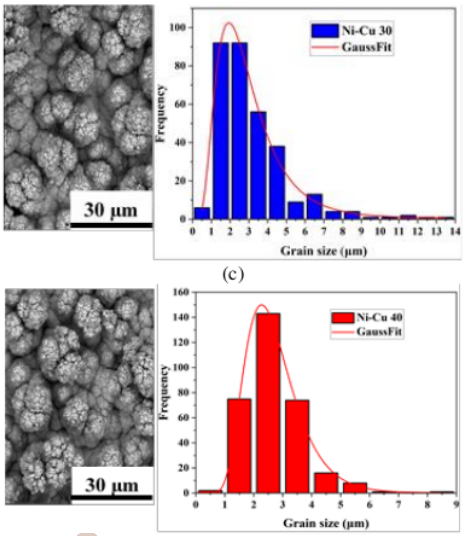
The surface morphological structure and the cross section of the coated samples were observed using SEM. Figure 6 shows SEM micrographs of the four samples' surface morphological structures. The deposits developed a fine-grained and compact spherical shape when lower deposition current densities were applied (Figure 5(a)). Deo et al. (27) and Goranova et al. (37) also observed this morphology at low current densities. The shape changed to a coarser cauliflower form when higher current densities were applied (Figure 6(d)). A diffusion-limited deposition mechanism in which a multigeneration spherical diffusion layer creates a cauliflower shape is likely to produce this type of morphology (40). As the current density increased, the cauliflower-like protrusions became more spaced and separated, creating gaps. The Ni-Cu 40 sample, produced with the highest current density, was found to have the largest gaps between the cauliflower-like bulges.

The increased nucleation rate can explain the observed decrease in grain size with increasing current density (41, 42). Ni-Cu ion flow to the cathode is faster at higher current densities. Further investigation is needed to determine the exact relationship between the current density and the grain size of the coating. In the Ni-Cu 15, Ni-Cu 20, Ni-Cu 30, and Ni-Cu 40 samples, the peak that corresponded to the (111) plane shifted toward the right (see Figure 2b). The reduction in d-spacing is ascribed to residual stress induced at a higher deposition rate (43). The statistical results of the grain size distribution presented in Figure 6 indicate that the grain size ranged from 4.63 to 1.94 µm. The decrease in grain size with the increase in current density is evident in the data shown in Table 5.



(a)

(b)



(c)

(d)

Figure 6. Surface SEM images of the deposited Ni-Cu alloy coatings and plots showing the statistical distribution of the grain size

Figure 7 depicts the relationship between the deposited alloy composition and the applied current density. EDS was used to determine the composition.

Figure 8 (a–d) displays SEM cross-section images of the produced Ni-Cu coatings. The absence of cracks between the substrate and coating demonstrates that appropriate adhesion occurred between the two entities (44). The thickness of the electrodeposited Ni-Cu coating was also measured for each sample (27), and the

TABLE 5. Average grain size found in each sample

Sample	Average grain size (µm)
Ni-Cu 15	4.63 ± 0.269
Ni-Cu 20	4.38 ± 0.365
Ni-Cu 30	2.28 ± 0.068
Ni-Cu 40	1.94 ± 0.032

following results were recorded: Ni-Cu 15 = 32 µm, Ni-Cu 20 = 42 µm, Ni-Cu 30 = 49 µm, and Ni-Cu 40 = 50 µm. The effect of the current density on the thickness of the Ni-Cu coating is depicted in Figure 8; the thickness increased as the current density increased (19). Hence, a higher current density results in more mass and a thicker coating. The findings presented in Figure 1 show that as the current density increased, so too did the CCE.

According to Faraday's law, when the deposition time remains constant for all samples, coatings formed at lower current densities will be thinner than those produced at higher current densities. A thinner covering may lead to severe interference from the Al substrate. The cathodic current density also affects the coating, as Deo et al. (27) discovered that increasing the current density improves the thickness of the film due to an increase in the CCE.

3. 4. Magnetic Properties Figure 9 depicts the magnetic characteristics and fluctuations in the current density magnetization measured with a VSM at room temperature (45). The results of the VSM analysis demonstrate that the coatings in the Ni-Cu 15, Ni-Cu 20, Ni-Cu 30, and Ni-Cu 40 samples displayed ferromagnetic activity. The low ferromagnetic activity of the $Cu_{54.95}Ni_{45.05}$ alloy film of Ni-Cu 15 could be attributed to Ni diffusion in the Cu matrix, as Cu is a diamagnetic metal and Ni is a ferromagnetic metal (46). The ferromagnetic properties of the Ni-Cu films of Ni-Cu 20, Ni-Cu 30, and Ni-Cu 40 increased with the Ni content of the alloy coatings.

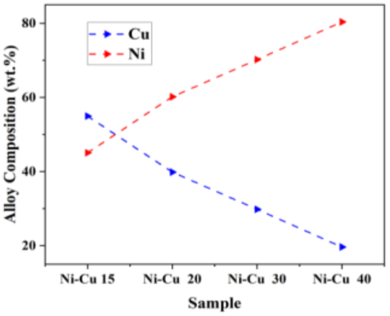


Figure 7. The dependence of the deposited alloy's composition (shown as wt% of the single electrolytes) on current density

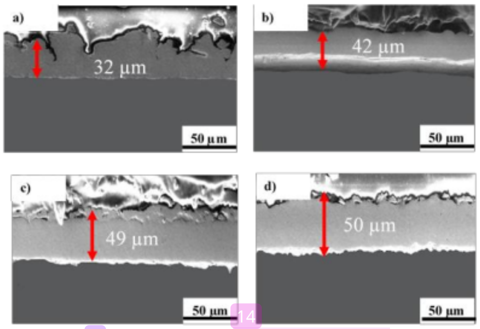
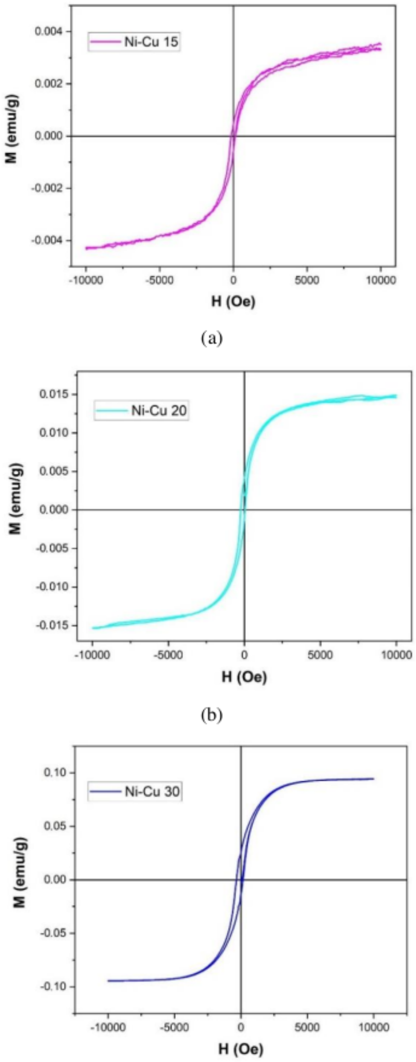


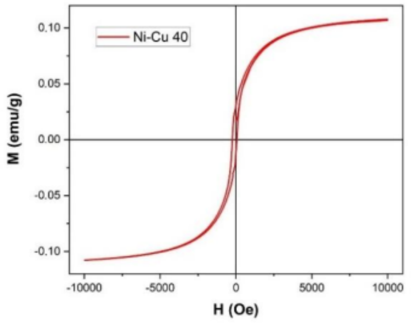
Figure 8. SEM cross-section images of the Ni-Cu coating on Al in the (a) Ni-Cu 15, (b) Ni-Cu 20, (c) Ni-Cu 30, and (d) Ni-Cu 40 samples



(a)

(b)

(c)



(d)

Figure 9. Hysteresis loops of multilayers generated at different current densities

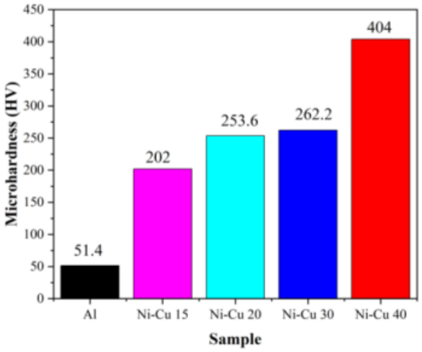


Figure 10. The microhardness of the electrodeposited Ni-Cu samples produced with different current densities

As the Ni content of the Ni-Cu alloy coatings increased, so did the saturation magnetization (see Table 6). Wang et al. (43) reported that saturation magnetization depends on the Ni content of Ni-Cu alloy coatings. In addition, Awasthi (22) reported that magnetization is enhanced by increasing the coating thickness. A possible reason for this enhanced magnetization is the magnetic disorder caused by the coating. It has been shown that the trend in saturation magnetization enhancement is associated with the coating level (47). Demidenko et al. (48) found that monel has paramagnetic properties at room temperature, while the Ni-Cu alloys in this study had ferromagnetic properties. Here, we have improved upon the results of previous studies in which phosphorus (P) and tungsten (W) were added (25). In this study, the Ni-Cu 40 sample exhibited the strongest magnetic properties.

3. 5. Hardness Figure 10 depicts the dependence of the microhardness of the Ni-Cu coatings on the current density in the plating bath. From the data presented in Figure 10, it is clear that the coating of the Ni-Cu 40 sample had the highest microhardness value (404 HV). In general, the microhardness increased with the current density and was attributed to the grain size and thickness of the coatings (42). Pingale et al. (19) found that hardness increased with the thickness of the coating. The coating thickness could influence the hardness of Ni-Cu films (26).

The results indicate that alloys with a greater Ni content are mechanically harder. The overall dependence of hardness and microhardness on the percentage of Ni is shown in Figure 11, and the data indicate that microhardness increases as the percentage of Ni increase. This result is similar to Marenych's (49) finding that the hardness value is highest with the highest Ni composition.

Moreover, the hardness reported in previous studies that resulted from electrodeposition of Ni-Cu on Al in the presence of P was lower than that recorded in the present study (24). This is due to the smaller grain size that resulted from applying a different current density. In addition, the increase in microhardness reported here is related to the role that Ni atoms play in grain refinement (31). Ramkumar et al. (50) reported a monel hardness value of 165 HV, which is lower than the peak hardness value recorded in the current study.

TABLE 6. The results of the magnetic analysis of the Ni-Cu/Al samples

Sample	Hc (Oe)	Mr (emu/g)	Ms (emu/g)
Ni-Cu 15	143.829	0.00047	0.003
Ni-Cu 20	140.081	0.004	0.015
Ni-Cu 30	256.215	0.025	0.094
Ni-Cu 40	144.023	0.032	0.108

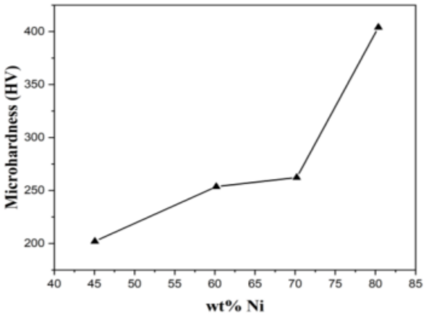


Figure 11. The relationship found between wt% Ni and microhardness, based on the data obtained from the four experimental samples

4. CONCLUSION

In this study, Ni-Cu alloys were electrodeposited onto Al substrates using citrate baths. The effects of adding a citrate solution on the properties of the deposited coatings were studied, and the results illustrate that single-phase Ni-Cu alloy layers were produced on the Al surface at all current densities. The CCE increased as the current density increased. The coatings formed at lower current densities showed a more compact and spherical morphology, while those formed at higher current densities showed a less uniform structure with a cauliflower-like morphology. Both the surface morphology and composition of the coating showed a strong dependence on the current density. The Ni-Cu alloy coating deposited at a low current density had a layer thickness of 32 µm, while the coating deposited at a high current density had a thickness of 50 µm. The saturation magnetization of the coating increased with the Ni content in the Ni-Cu alloy and with the coating thickness. The hardness increased with the coating thickness, grain size, and Ni composition in the solution. The hardness of the produced Ni-Cu alloy coatings was found to be greater than that of monel.

5. ACKNOWLEDGMENTS

The authors extend their gratitude to the Ministry of Research, Technology and Higher Education of the Republic of Indonesia for the financial support of Hibah Penelitian Disertasi Doktor No: NKB-971/UN2.RST/HKP.05.00/2022.

6. REFERENCES

1. Chen Z, Wang C, Tang C, Lek YZ, Kandukuri SY, Du H, et al. Microstructure and mechanical properties of a Monel K-500 alloy fabricated by directed energy deposition. Materials Science and Engineering: A. 2022;857:144113. https://doi.org/10.1016/j.msea.2022.144113
2. Kukliński M, Bartkowska, A., and Przestacki, D. . Microstructure and selected properties of Monel 400 alloy after laser heat treatment and laser boriding using diode laser. International Journal of Advanced Manufacturing Technology. 2018;98(9-12): 3005–17. https://doi.org/10.1007/s00170-018-2343-9
3. Mohagheghpour E, Larijani M, Rajabi M, Gholamipour R. Effect of Silver Clusters Deposition on Wettability and Optical Properties of Diamond-like Carbon Films. International Journal of Engineering, Transactions C: Aspects, . 2021;34(3):706-13. https://doi.org/10.5829/ije.2021.34.03c.15
4. Nady H, Negem M. Ni–Cu nano-crystalline alloys for efficient electrochemical hydrogen production in acid water. RSC advances. 2016;6(56):51111-9. https://doi.org/10.1039/c6ra08348j
5. Negem M, Nady H. Electroplated Ni-Cu nanocrystalline alloys and their electrocatalytic activity for hydrogen generation using alkaline solutions. international journal of hydrogen energy. 2017;42(47):28386-96. https://doi.org/10.1016/j.ijhydene.2017.09.147
6. Guisbiers G, Khanal S, Ruiz-Zepeda F, De La Puente JR, José-Yacaman M. Cu–Ni nano-alloy: mixed, core–shell or Janus nano-particle? Nanoscale. 2014;6(24):14630-5. https://doi.org/10.1039/c4nr05739b
7. Mae Y. What the Darken–Gurry plot means about the solubility of elements in metals. Metallurgical and Materials Transactions A. 2016;47(12):6498-506. https://doi.org/10.1007/s11661-016-3730-1
8. Alizadeh M, Safaei H. Characterization of Ni-Cu matrix, Al2O3 reinforced nano-composite coatings prepared by electrodeposition. Applied Surface Science. 2018;456:195-203. https://doi.org/10.1016/j.apsusc.2018.06.095
9. Toghraei M, Siadati H. Electrodeposited co-pi catalyst on α-Fe2O3 photoanode for water-splitting applications. International Journal of Engineering. 2018;31(12):2085-91. 10.5829/ije.2018.31.12c.13
10. KK P. Experimental investigation by cryogenic treatment of aluminium 6063 and 8011 and nicow coating to improve hardness and wear. International Journal of Engineering. 2016;29(6):827-33. https://doi.org/10.5829/idosi.ije.2016.29.06c.12
11. Kumar D, Angra S, Singh S. Mechanical properties and wear behaviour of stir cast aluminum metal matrix composite: a review. International Journal of Engineering. 2022;35(4):794-801. https://doi.org/10.5829/IJE.2022.35.04A.19
12. Moosaei H, Zareei A, Salemi N. Elevated Temperature Performance of Concrete Reinforced with Steel, Glass, and Polypropylene Fibers and Fire-proofed with Coating. International Journal of Engineering. 2022;35(5):917-30. https://doi.org/10.5829/ije.2022.35.05b.08
13. Poursaeidi E, Salarvand A. Comparison of properties of ti/tin/ticn/tialn film deposited by cathodic arc physical vapor and plasma-assisted chemical vapor deposition on custom 450 steel substrates. International Journal of Engineering. 2016;29(10):1459-68. https://doi.org/10.5829/idosi.ije.2016.29.10a.17
14. Allahyarzadeh M, Aliofkhazraei M, Rouhaghdam AS, Torabinejad V. Gradient electrodeposition of Ni-Cu-W (alumina) nanocomposite coating. Materials & Design. 2016;107:74-81. https://doi.org/10.1016/j.matdes.2016.06.019
15. Geramipour F, Khoei SM, Gugtapeh HS. Effect of shaped waveform on structure and electrochemical corrosion behavior of pulse electrodeposited NiCu alloy coatings. Surface and Coatings Technology. 2021;424:127643. https://doi.org/10.1016/j.surfcoat.2021.127643
16. Thurber CR, Ahmad YH, Sanders SF, Al-Shenawa A, D'Souza N, Mohamed AM, et al. Electrodeposition of 70-30 Cu–Ni nanocomposite coatings for enhanced mechanical and corrosion properties. Current Applied Physics. 2016;16(3):387-96. https://doi.org/10.1016/j.cap.2015.12.022
17. Goranova D, Rashkov R, Avdeev G, Tonchev V. Electrodeposition of Ni–Cu alloys at high current densities: details of the elements distribution. Journal of Materials Science. 2016;51:8663-73. https://doi.org/10.1007/s10853-016-0126-y
18. Lee W-H, Chung K. Investigation of a copper–nickel alloy resistor using co-electrodeposition. Journal of Applied Electrochemistry. 2020;50:535-47. https://doi.org/10.1007/s10800-020-01398-0
19. Pingale AD, Belgamwar SU, Rathore JS. Effect of graphene nanoplatelets addition on the mechanical, tribological and corrosion properties of Cu–Ni/Gr nanocomposite coatings by electro-co-deposition method. Transactions of the Indian Institute of Metals. 2020;73:99-107. https://doi.org/10.1007/s12666-019-01807-9

20. Hughes RA, Menumerov E, Neretina S. When lithography meets self-assembly: a review of recent advances in the directed assembly of complex metal nanostructures on planar and textured surfaces. Nanotechnology. 2017;28(28):282002. https://doi.org/10.1088/1361-6528/aa77ce
21. Kamel M, Anwer Z, Abdel-Salam I, Ibrahim I. Electrodeposition of nanocrystalline Ni–Cu alloy from environmentally friendly lactate bath. Surface and Interface Analysis. 2014;46(7):442-8. https://doi.org/10.1002/sia.5525
22. Awasthi S, Pandey SK, Balani K. Tuning the magnetism and tribological behaviour of electrodeposited Ni/Cu bi-layer by selective reinforcement of carbon nanotubes. Journal of Alloys and Compounds. 2020;818:153287. https://doi.org/10.1016/j.jallcom.2019.153287
23. Heidarzadeh A, Saeid T. Correlation between process parameters, grain size and hardness of friction-stir-welded Cu–Zn alloys. Rare Metals. 2018;37:388-98. https://doi.org/10.1007/s12598-016-0704-9
24. Karunakaran M, Pugazhvadivu M, Gunasegaran V, Gowtham G. Electrodeposition of Cu-Ni-PW Composite on Al-6063 Substrate. 2018. https://doi.org/10.26438/ijsrpas/v6i3.5964
25. Karunakaran M, Vadivu MP. Magnetic and micro-mechanical behavior of Cu-Ni-PW-TiO2 hybrid composite electroplating on Al alloy substrate. Journal of Magnetism and Magnetic Materials. 2019;475:359-67. https://doi.org/10.1016/j.jmmm.2018.11.077
26. Soegijono B, Susetyo F, editors. Magnetic field exposure on electroplating process of ferromagnetic nickel ion on copper substrate. Journal of Physics: Conference Series; 2022: IOP Publishing.
27. Deo Y, Guha S, Sarkar K, Mohanta P, Pradhan D, Mondal A. Electrodeposited Ni-Cu alloy coatings on mild steel for enhanced corrosion properties. Applied Surface Science. 2020;515:146078. https://doi.org/10.1016/j.apsusc.2020.146078
28. Seakr R. Microstructure and crystallographic characteristics of nanocrystalline copper prepared from acetate solutions by electrodeposition technique. Transactions of Nonferrous Metals Society of China. 2017;27(6):1423-30. https://doi.org/10.1016/S1003-6326(17)60164-X
29. Dolabella S, Borzì A, Dommann A, Neels A. Lattice strain and defects analysis in nanostructured semiconductor materials and devices by high-resolution X-ray diffraction: theoretical and practical aspects. Small Methods. 2022;6(2):2100932. https://doi.org/10.1002/smtd.202100932
30. Budi S, Tawwabin RA, Cahyana U, Paristiowati M. Saccharin-assisted galvanostatic electrodeposition of nanocrystalline FeCo films on a flexible substrate. International Journal of Electrochemical Science. 2020;15(7):6682-94. https://doi.org/10.20964/2020.07.74
31. Syamsuir S, Soegijono B, Yudanto SD, Basori B, Ajiriyanto MK, Nanto D, et al. Electrolyte Temperature Dependency of Electrodeposited Nickel in Sulfate Solution on the Hardness and Corrosion Behaviors. International Journal of Engineering, Transactions C: Aspects, . 2023;36(6):1193-200. 10.5829/ije.2023.36.06c.18
32. Soegijono B, Susetyo FB, Fajrah MC. Electrodeposition of Paramagnetic Copper Film under Magnetic Field on Paramagnetic Aluminum Alloy Substrates. e-Journal of Surface Science and Nanotechnology. 2020;18:281-8. https://doi.org/10.1380/EJSSNT.2020.281
33. Devi C, Ashokkumar R. INFLUENCE OF DIFFERENT CURRENT DENSITY ON CHARACTERISTICS OF NiFeP NANO ALLOY THIN FILMS. Rasayan Journal of Chemistry. 2018;11(3). https://doi.org/10.31788/RJC.2018.1133088
34. Li B, Mei T, Li D, Du S. Ultrasonic-assisted electrodeposition of Ni-Cu/TiN composite coating from sulphate-citrate bath: Structural and electrochemical properties. Ultrasonics sonochemistry. 2019;58:104680. https://doi.org/10.1016/j.ultsonch.2019.104680
35. Sarac U, Baykul MC. Morphological and microstructural properties of two-phase Ni–Cu films electrodeposited at different electrolyte temperatures. Journal of alloys and compounds. 2013;552:195-201. https://doi.org/10.1016/j.jallcom.2012.10.071
36. Arasteh J. Microhardness Optimization of Al–TiC Nanocomposite Produced by Mechanical Milling and Heat Treatment. Advanced Ceramics Progress. 2021;7(1):35-45. https://doi.org/10.30501/ACP.2021.265197.1052
37. Goranova D, Avdeev G, Rashkov R. Electrodeposition and characterization of Ni–Cu alloys. Surface and Coatings Technology. 2014;240:204-10. https://doi.org/10.1016/j.surfcoat.2013.12.014
38. Hu G, Huang R, Wang H, Zhao Q, Zhang X. Facile galvanic replacement deposition of nickel on copper substrate in deep eutectic solvent and its activation ability for electroless Ni–P plating. Journal of Solid State Electrochemistry. 2022;26(5):1313-22. https://doi.org/10.1007/s10008-022-05172-4
39. Niu J, Song M, Zhang Y, Zhang Z. Dealloying induced nanoporosity evolution of less noble metals in Mg ion batteries. Journal of Magnesium and Alloys. 2021;9(6):2122-32. https://doi.org/10.1016/j.jma.2021.04.003
40. Wang S, Guo X, Yang H, Dai J, Zhu R, Gong J, et al. Electrodeposition mechanism and characterization of Ni–Cu alloy coatings from a eutectic-based ionic liquid. Applied Surface Science. 2014;288:530-6. https://doi.org/10.1016/j.apsusc.2013.10.065
41. Nwaeju CC, Eboh, A. O., and Edoziuno, F. O. . Grain size evolution mechanical and corrosion behaviour of precipitate strengthened Cu-Ni alloy. Acta Metallurgica Slovaca. 2022;28(4):188–96. https://doi.org/10.36547/ams.28.4.1609
42. Ameri Ekhtiarabadi T, Zandrahimi M, Ebrahimifar H. The Impact of Current Density of Electroplating on Microstructure and Mechanical Properties of Ni-ZrO2-TiO2 Composite Coating. Advanced Ceramics Progress. 2020;6(1):22-9. https://doi.org/10.30501/acp.2020.233518.1038
43. Wang C, Hossain Bhuiyan ME, Moreno S, Minary-Jolandan M. Direct-write printing copper–nickel (Cu/Ni) alloy with controlled composition from a single electrolyte using co-electrodeposition. ACS applied materials & interfaces. 2020;12(16):18683-91. https://doi.org/10.1021/acsami.0c01100
44. Hanachi M, Seyedraoufi Z, Abouei V. Investigation of Microstructure, Hardness, and Corrosion Resistance of Ni-P-GO Electroless Nanocomposite Coating on AZ31D Alloy Surface. Advanced Ceramics Progress. 2020;6(3):55-62. https://doi.org/10.30501/acp.2020.233518.1038
45. Dhara B, Jha PK, Gupta K, Bind VK, Ballav N. Diamagnetic Molecules Exhibiting Room-Temperature Ferromagnetism in Supramolecular Aggregates. The Journal of Physical Chemistry C. 2017;121(22):12159-67. https://doi.org/10.1021/acs.jpcc.7b02145
46. Qasim I, Waqee-ur-Rehman M, Mumtaz M, Hussain G, Nadeem K, Shehzad K. Ferromagnetic (Ni) nanoparticles–CuTl-1223 superconductor composites. Journal of Magnetism and Magnetic Materials. 2016;403:60-7. https://doi.org/10.1016/j.jmmm.2015.11.066
47. Padmapriya G, Manikandan A, Krishnasamy V, Jaganathan SK, Antony SA. Enhanced catalytic activity and magnetic properties of spinel Mn x Zn 1− x Fe 2 O 4 (0.0≤ x≤ 1.0) nano-photocatalysts by microwave irradiation route. Journal of Superconductivity and Novel Magnetism. 2016;29:2141-9. https://doi.org/10.1007/s10948-016-3527-x

48. Demidenko O, Zhyvulka A, Yanushkevich K, Galias A, Constantin V, Neacsu E, et al. Magnetic properties of stainless steels under corrosive action of based on choline chloride ionic liquids. Journal of Magnetism and Magnetic Materials. 2019;477:74-6. https://doi.org/10.1016/j.jmmm.2019.01.034

49. Marenych O, Ding D, Pan Z, Kostryzhev A, Li H, van Duin S. Effect of chemical composition on microstructure, strength and wear resistance of wire deposited Ni-Cu alloys. Additive Manufacturing. 2018;24:30-6. https://doi.org/10.1016/j.addma.2018.08.003

50. Ramkumar KD, Joshi V, Pandit S, Agrawal M, Kumar OS, Periwal S, et al. Investigations on the microstructure and mechanical properties of multi-pass pulsed current gas tungsten arc weldments of Monel 400 and Hastelloy C276. Materials & Design. 2014;64:775-82. https://doi.org/10.1016/j.matdes.2014.08.055

Persian Abstract

چکیده

پوشش‌های آلیاژی تک فاز نیکل مسغنی از نیکل (Ni) بر روی بسترهای آلومینیومی (Al) با رسوب الکتریکی در حمام سیترات تثبیت شده تولید شدند. آزمایش‌های رسوب الکتریکی در چهار چگالی جریان مختلف انجام شد. افزایش چگالی جریان منجر به افزایش سرعت رسوب فلز سریعتر از نرخ تکامل هیدروژن شد. بنابراین، راندمان جریان کاتدی افزایش یافت. سیستم‌های کریستالی آلیاژهای Ni-Cu مکعبی در مرکز سطح (FCC) بودند، با صفحه (۱۱۱) به عنوان صفحه کریستالی ترجیحی. میکروسکوپ الکترونی روبشی با اندازه‌گیری‌های طیف‌سنجی پرتو ایکس پراکنده انرژی (SEM-EDS) نشان داد که محتوای نیکل در پوشش با افزایش چگالی جریان افزایش می‌یابد. نمونه Ni-Cu 40 بیشترین مقدار نیکل را داشت و مورفولوژی همگن و فشرده را نشان داد. مشخص شد که هر چه غلظت نیکل در محلول بیشتر باشد، اندازه دانه کوچکتر است. اندازه‌گیری‌های ثبت‌شده با یک مغناطیس‌سنج نمونه ارتعاشی (VSM) نشان داد که نمونه Ni-Cu 40 اشباع مغناطیسی را ارائه می‌کند، با بالاترین مقدار ۰.۱۰۸ emu/g. روش میکروسختی ٤٠٤ HV بر روی نمونه Ni-Cu 40 تولید کرد. در نتیجه، چگالی جریان بالاتر با ترکیب نیکل بالاتر و افزایش ضخامت همراه بود که مسئول افزایش خواص مغناطیسی و سختی بود.

Effect of Current Density on Magnetic and Hardness Properties of Ni-Cu Alloy Coated on Al via Electrodeposition

ORIGINALITY REPORT

17%	7%	17%	1%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Yashwardhan Deo, Sounak Guha, Kuntal Sarkar, Puspanjali Mohanta, Debabrata Pradhan, Avik Mondal. "Electrodeposited Ni-Cu alloy coatings on mild steel for enhanced corrosion properties", Applied Surface Science, 2020 Publication	3%
2	Norman B. Pilling, Robert E. Bedworth. "Oxidation of Copper-Nickel Alloys at High Temperatures.", Industrial & Engineering Chemistry, 1925 Publication	2%
3	Shikha Awasthi, Sarvesh Kumar Pandey, Kantesh Balani. "Tuning the magnetism and tribological behaviour of electrodeposited Ni/Cu bi-layer by selective reinforcement of carbon nanotubes", Journal of Alloys and Compounds, 2020 Publication	1%

4	Fatih Doğan, Mehmet Uysal, Erhan Duru, Hatem Akbulut, Serdar Aslan. "Pulsed electrodeposition of Ni-B/TiN composites: effect of current density on the structure, mechanical, tribological, and corrosion properties", Journal of Asian Ceramic Societies, 2020 Publication	1 %
5	www.mdpi.com Internet Source	1 %
6	R. SEAKR. "Microstructure and crystallographic characteristics of nanocrystalline copper prepared from acetate solutions by electrodeposition technique", Transactions of Nonferrous Metals Society of China, 2017 Publication	1 %
7	Morteza Alizadeh, Hamed Safaei. "Characterization of Ni-Cu matrix, Al 2 O 3 reinforced nano-composite coatings prepared by electrodeposition", Applied Surface Science, 2018 Publication	1 %
8	Muthusankar Ganesan, Chao-Chen Liu, Sabarison Pandiyarajan, Chen-Ta Lee, Ho-Chiao Chuang. "Post-supercritical CO2 electrodeposition approach for Ni-Cu alloy fabrication: An innovative eco-friendly	1 %

strategy for high-performance corrosion resistance with durability", Applied Surface Science, 2021
Publication

9	Weiwei Zhang, Wenzhe Xia, Baosong Li, Mingyuan Li, Ming Hong, Zhen Zhang. "Influences of Co and process parameters on structure and corrosion properties of nanocrystalline Ni-W-Co ternary alloy film fabricated by electrodeposition at low current density", Surface and Coatings Technology, 2022 Publication	1%
10	nagaokaut.repo.nii.ac.jp Internet Source	1%
11	www.scielo.br Internet Source	1%
12	Begum Unveroglu. "Electrodeposition and Characterization of Ni-Cu Alloy and Submicron-Sized CeO2 Reinforced Ni-Cu Metal Matrix Composite Coatings", Arabian Journal for Science and Engineering, 2022 Publication	<1%
13	Haifeng Tan, Wenchao Yang, Mingzhu Hao, Chao Wang, Jie Yang, Haixuan Sunyu, Yunhe Ling, Guihong Song, Chunlin He. "Effect of Current Density on the Corrosion Resistance and Photocatalytic Properties of Cu-Ni-	<1%

Zn0.96Ni0.02Cu0.02O Nanocomposite Coatings", Materials, 2023
Publication

---

14 Pengyuan Zhang, Guozhe Meng, Yanqiu Wang, Bing Lei, Fuhui Wang. "Significantly enhanced corrosion resistance of Ni–Cu coating modified by minor cerium", Corrosion Communications, 2021 <1 %
Publication

---

15 japsonline.com <1 %
Internet Source

---

16 M. Karunakaran, M. Pugazh Vadivu. "Magnetic and micro-mechanical behavior of Cu-Ni-P-W-TiO2 hybrid composite electroplating on Al alloy substrate", Journal of Magnetism and Magnetic Materials, 2018 <1 %
Publication

---

17 Sarac, Umut, and M. Celalettin Baykul. "Morphological and microstructural properties of two-phase Ni–Cu films electrodeposited at different electrolyte temperatures", Journal of Alloys and Compounds, 2013. <1 %
Publication

---

18 F. Geramipour, S.M. Mousavi Khoei, H. Shooshtari Gugtapeh. "Effect of shaped waveform on structure and electrochemical corrosion behavior of pulse electrodeposited <1 %

Ni Cu alloy coatings", Surface and Coatings Technology, 2021
Publication

---

19 www.acerp.ir
Internet Source
<1%

---

20 Kyle D. Gilroy, Aleksey Ruditskiy, Hsin-Chieh Peng, Dong Qin, Younan Xia. "Bimetallic Nanocrystals: Syntheses, Properties, and Applications", Chemical Reviews, 2016
Publication
<1%

---

21 Shaohua Wang, Xingwu Guo, Haiyan Yang, JiChun Dai, Rongyu Zhu, Jia Gong, Liming Peng, Wenjiang Ding. "Electrodeposition mechanism and characterization of Ni–Cu alloy coatings from a eutectic-based ionic liquid", Applied Surface Science, 2014
Publication
<1%

---

22 Yuan Yuan, Yuanwei Chen, Jia-Hui Liu, Haifang Wang, Yuanfang Liu. "Biodistribution and fate of nanodiamonds in vivo", Diamond and Related Materials, 2009
Publication
<1%

---

23 Submitted to Birla Institute of Technology and Science Pilani
Student Paper
<1%

---

24 P. Gupta. "Effect of annealing current density on the microstructure of nanocrystalline
<1%

FeCuNbSiB alloy", Journal of Applied Physics, 2007
Publication

---

25 Baosong Li, Tianyong Mei, Dandan Li, Shengsong Du. "Ultrasonic-assisted electrodeposition of Ni-Cu/TiN composite coating from sulphate-citrate bath: Structural and electrochemical properties", Ultrasonics Sonochemistry, 2019 <1%
Publication

---

26 D. Goranova, R. Rashkov, G. Avdeev, V. Tonchev. "Electrodeposition of Ni–Cu alloys at high current densities: details of the elements distribution", Journal of Materials Science, 2016 <1%
Publication

---

27 Torabinejad, V., M. Aliofkhazraei, S. Assareh, M.H. Allahyarzadeh, and A. Sabour Rouhaghdam. "Electrodeposition of Ni-Fe alloys, composites, and nano coatings–A review", Journal of Alloys and Compounds, 2017. <1%
Publication

---

28 Zuveria Firdouz, Pragya Tripathi, K. Mondal, Kantesh Balani. "Effect of carbonaceous reinforcements on anticorrosive and magnetic properties of Ni Cu based composite coatings prepared by pulsed <1%

electrodeposition", Surface and Coatings Technology, 2022
Publication

Exclude quotes	On	Exclude matches	< 15 words
Exclude bibliography	On		

Effect of Current Density on Magnetic and Hardness Properties of Ni-Cu Alloy Coated on Al via Electrodeposition

GRADEMARK REPORT

FINAL GRADE	GENERAL COMMENTS
/0	

PAGE 1

PAGE 2

PAGE 3

PAGE 4

PAGE 5

PAGE 6

PAGE 7

PAGE 8

PAGE 9

PAGE 10

PAGE 11