I. R. IRAN

ISSN: 2423-7167
e-ISSN: 1735-9244

International Journal of Engineering

Journal Homepage: www.ije.ir

MERC

TRANSACTIONS C: Aspects

Volume 34, Number 09, September2021

Materials and Energy Research Center

In the Name of God

INTERNATIONAL JOURNAL OF ENGINEERING
Transactions C: Aspects

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)

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 6 journal pages (including figures and tables). For the additional to 6 journal pages, there will be page charges. It is advised to keep papers limited to 3500 words.

Page Charges for Papers More Than 6 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

ISSN: 2423-7167, e-ISSN: 1735-9244

International Journal of Engineering, Volume 34, Number 9, September 2021

CONTENTS:

ISSN: 2423-7167, e-ISSN: 1735-9244

International Journal of Engineering, Volume 34, Number 9, September 2021

International Journal of Engineering

Journal Homepage: www.ije.ir

An Improvement on Slot Configuration Structure for a Low-speed Surface-mounted Permanent Magnet Synchronous Generator with a Wound Cable Winding

H. Parivar, S. M. Seyyedbarzegar*, A. Darabi

Department of Electrical Engineering, Shahrood University of Technology, Shahrood, Iran

PAPER INFO

Paper history:
Received 13 March 2021
Received in revised form 07 May 2021
Accepted 15 June 2021

Keywords:
Marine Energy Conversion System
Tidal Energy Conversion System
Cross-linked Polyethylene
Finite-element Method
Torque Ripple
Joule Loss

ABSTRACT

This paper concentrates on low-speed surface-mounted permanent magnet synchronous generator (LS-SMPMSG) electromagnetic design using finite-element (FEM)-based method. The introduced LS-SMPMSG is utilized in marine energy conversion (MEC) and tidal energy conversion (TEC) systems. The main feature of the designed LS-SMPMSG is its ability to generate electrical power in MEC and TEC systems at a low speed of approximately 100 rpm. In the second part, a new slot structure for LS-SMPMSG has been proposed to enhance the LS-SMPMSG performance. The FEM analysis shows that the new slot structure led to reducing the torque ripple and the cogging torque equal to 19.54% and 87.5% respectively. Also, it influenced the LS-SMPMSG Joule losses by 16.1% reduction. Another advantage of the new slot structure of LS-SMPMSG is its easy construction and assembly. The FEM analysis validates the advantages of the proposed structure.

***doi**: 10.5829/ije.2021.34.09c.01*

1. INTRODUCTION

Permanent magnet (PM) machines are used on a large scale, remarkably in the past two decades [1, 2]. Two important factors that have been highlighted in the utilization of PM machines are cost competition and the availability of rare-earth PMs. Throughout the early ten years, relying on vast resources of rare PM materials, China's production began to flourish. Due to these immense resources, the cost of the PM diminished and made it economical to use [3]. Meaningfully, the electrical machine efficacy has been enhanced by introducing and utilizing PMs [4]. Thus, in the last 20 years, not only the cost of magnets has decreased, but also various categorizations of magnets became available [5]. This has made it conceivable to utilize PM machines even in industries suchlike the aircraft industry or the military. In the present circumstances, PMs with a BH_{max} of more than 60 *MG.Oe* are manufactured and marketed by assorted companies [6].

*Corresponding Author Institutional Email: seyyedbarzegar@shahroodut.ac.ir (S. M. Seyyedbarzegar)

With advances in efficiency-oriented applications, the prominence of economical use of resources has originated the need for compact and high-efficiency electromechanical systems. In practical terms, it conveys that conventional electric machines are being substituted by new models. One of the electrical machines which produce electricity from renewable energy resources is PM machines [7]. Owing to the high-power density of PM material, PM machines might have an upper power density and high efficiency through the absence of excitation, magnetization currents, and insignificant rotor losses. Undoubtedly, PM machines play an indispensable assignment in energy production. Ocean energy is one such emerging energies that have originated promising developments in the energy industry. Momentous commercial developments have been initiated in the US, Europe, Asia, and South America over the next 10 to 20 years [8]. Plenty of generators can be utilized in the wave energy conversion (WEC) systems suchlike squirrel cage induction generator (SCIG), synchronous generator (SG), PM synchronous generator (PMSG). Table 1, illustrated the speed extent of the ordinary generator

Please cite this article as: H. Parivar, S. M. Seyyedbarzegar, A. Darabi, An Improvement on Slot Configuration Structure of a Low-speed Surface-mounted Permanent Magnet Synchronous Generator with a Wound Cable Winding, International Journal of Engineering, Transactions C: Aspects Vol. 34, No. 9, (2021) 2045-2052

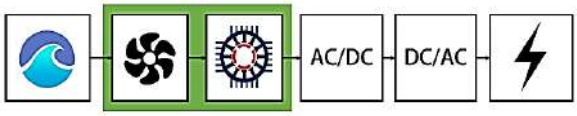


Figure 1. The Wave Energy Conversion (WEC) system

TABLE 1. Generator speed ranges for each WEC category [9]

Speed Range	Generator Type
0-400^{rpm}	Linear: PMSG, Rotary: Gearbox coupled, Variable speed SG / PMSG
100-250^{rpm}	Gearbox coupled, fixed speed SG / SCIG
600-1500^{rpm}	Variable speed PMSG/SCIG, Low-speed Variable-speed PMSG
1000-300^{rpm}	Variable-speed SCIG/SG/PMSG For low-storage designs: Variable-speed SG/PMSG/SCIG

types exploited in WEC systems [9]. Energy converters play a fundamental role in MEC and TEC systems. Electrical machines of the WEC branch can be classified in two subclasses: the rotary and the linear converter. PM machines are one of the brand-new rotary types which benefit from generating higher power due to the utilization of the PMs material [10]. Multiple factors impact the designating generators for MEC and TEC systems. For instance, economic and ecological aspects, generators and their drive characteristics, turbine technology, etc. Tidal currents are an exceedingly reputable source of renewable energy [11]. Customarily, convenience sites are the Bay of Fundy, Canada [12] and EMEC the Fall of Warness, UK tidal test site. Societies tend to take the advantage of renewable energy resources to solve the long-standing issues of fossil fuels concerning the rise in environmental sensitiveness [13]. A comprehensive WEC system is shown in Figure 1. At slow speeds, and by accessing the multitude of poles, on one hand, SPM benefits from an outstanding profit in terms of power density. On the other hand, the SPM generators also benefit from lower overall losses in comparison to IPM [14]. The neodymium ferrite (Nd_2Fe_{14}) corresponding to Figure 2 (a) in the proposed LS-SMPMSG, reduces the external diameter of the LS-SMPMSG which eventuates to a higher energy density, which is acceptable at low speeds. The hindrances in the utilization of Nd_2Fe_{14} could be stated as high operating temperature and considerable cost. The SPM owns higher output power, lower THD, uncomplicated and economical design, low iron loss, and ultimately, low magnet leakage flux. Accordingly, the SM-type is preferred to the LS-SMPMSG rotor. Figure 2 (b) represents the stator of the same LS-SMPMSG.

In this paper, a novel structure is proposed to design an LS-SMPMSG to be used in TEC and MEC systems. The simulation and analysis of the LS-SMPMSG have been conducted in JMAG.

2. FEM MODELLING OF LS-SMPMSG

2. 1. LS-SMPMSG Design and Assumptions In the framework of this study, a 7.2^{kW}, 100^{rpm} LS-SMPMSG has been designed. The stator is wounded by an XLPE cable with a conductor cross-section of 8 mm^2. Two layers of air and XLPE insulation, which is robust and vigorous insulation, are used to prevent any sparks in the generator. XLPE cable and the LS-SMPMSG parameters are shown in Table 2 and Table 3, respectively.

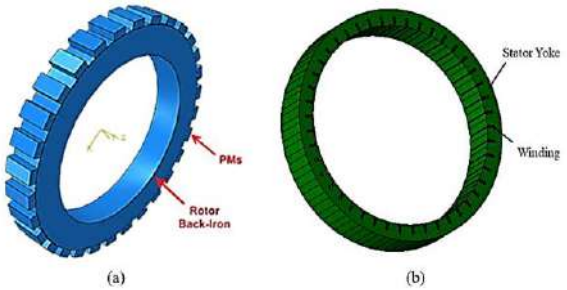


Figure 2. The LS-SMPMSG (a) rotor and (b)stator

TABLE 2. Typical properties of XLPE

Parameters	Volume
Density, g.cm^3	0.198
Tensile Strength, MPa	0.184
Elongation, %	500
Heat Distortion, %	20
Thermal Conductivity, w/m^{o}C at 90^{o}C	0.27
Modulus Elasticity, MPa	121
Dielectric Constant	2.3
Volume Resistivity, Ωm at 23°C	10^{16}

TABLE 3. LS-SMPMSG parameters

Parameters	Volume
Rated Power, KW	7.2
Speed, rpm	100
Coil Resistance, Ω	1.76
Stator Inner Radius, mm	220
Stator Outer Radius, mm	270
PM, mm	27.47×8
Cable Radius, mm	2
XLPE Insulation, mm	3.5
Air Insulation, mm	3.8
Slots/pole,-	9/3

2. 2. The LS-SMPMSG Modelling The electromagnetic field of the LS-SMPMSG is assumed as a two-dimensional problem. By combining group equations of the field and circuit, the LS-SMPMSG field equation to mode the time-dependent electrical fields are defined and described using Equation (1) [15].

$$\kappa\frac{\partial A_z}{\partial t}+\nabla\cdot\left(\nabla A_z\frac{1}{\mu_0\mu_r}\right)=-\kappa.E \tag{1}$$

where κ is electrical conductivity and μ_0, μ_r, A_z and E are the vacuum and relative permeability, the axial magnetic vector potential and applied field, respectively. In the circuit equation of the LS-SMPMSG (2)-(5), the coil end windings can be represented by coil end impedance. I_a, I_b and I_c are the conductor currents of three-phase *a*, *b* and *c*. U_a, U_b and U_c are three-phase terminal voltage obtained from solving field Eequation (1) and U_{ab}, U_{bc} and U_{cb} are the terminal line voltage between *ab*, *bc* and *ca* phases. *Rs (end)* and L_s*(end)* are coil end resistance and coil end inductance, respectively.

$$I_a+I_b+I_c=0 \tag{2}$$

$$U_{ab}=U_a+R_s(end)I_a+L_s(end)\frac{\partial I_a}{\partial t}-U_b -R_s(end)I_b-L_s(end)\frac{\partial I_b}{\partial t} \tag{3}$$

$$U_{bc}=U_b+R_s(end)I_b+L_s(end)\frac{\partial I_b}{\partial t}-U_c -R_s(end)I_c-L_s(end)\frac{\partial I_c}{\partial t} \tag{4}$$

$$U_{ca}=U_c+R_s(end)I_c+L_s(end)\frac{\partial I_c}{\partial t}-U_a -R_s(end)I_a-L_s(end)\frac{\partial I_a}{\partial t} \tag{5}$$

2. 3. The Material Used in LS-SMPMSG The materials utilized in the LS-SMPMSG are presented in Table 4. Under the appropriate operation of XLPE, the preferment of this research is to take advantages of the aforementioned insulation in the above-discussed LS-SMPMSG. In LS-SMPMSG stator used round cable winding which provides a three-phase magnetic field. Also, ArcelorMittal M800-100A and neodymium ferrite (Nd2Fe14) are used in the rotor, stator, and PM, respectively. The density of NdFeB is 7800 *kg/m*^{3}, thermal conductivity and specific heat are 7.6 *W/m/ ˚C* and 460 *J/kg/ ˚C*. The winding is a three-phase, 2 layers, and delta-connected. The range of turns in each phase is 35 and the phase resistance is *1.76* Ω. A non-uniform air gap can increase non-uniform forces in the air gap, which could lead to vibration subsequently irreversible mechanical damage to the synchronous machines. Also, based on the effect of a non-uniform air gap, the PM rotor's field is uncontrollable. Hence, a 4.5^{mm} air gap amidst the rotor and stator was chosen to avoid the non-uniform forces in the air gap and mechanical stresses. Another argument for choosing a large air gap is the ease of construction and assembly.

2. 4. Investigation of LS-SMPMSG Operation Conditions In this subsection, results of 2D analysis of the analytically designed LS-SMPMSG have been depicted. Figure 3 illustrates the LS-SMPMSG flux lines at no-load operation. Moreover Figure 4 shows the 2D flux density distribution in designed LS-SMPMSG. Figure 5 and Figure 6 indicate the Back-EMF and the three-phase stator current in the designed machine, accordingly.

TABLE 4. The material used in LS-SMPMSG

Parameters	Volume
Stator & Rotor Yoke	ArcelorMittal M800-100A
PM	NdFeB Hitachi Metals HIDENSE-500
Winding	Copper
Insulation. 1	XLPE
Insulation. 2	Air

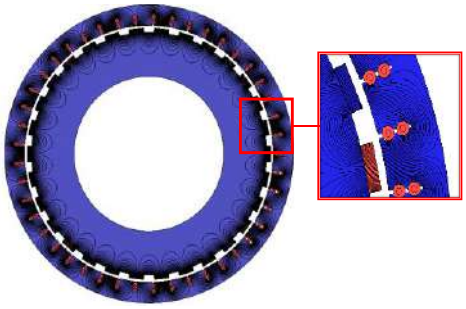


Figure 3. Magnetic flux lines of LS-SMPMSG

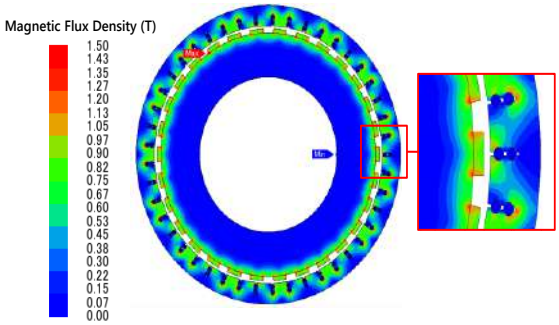


Figure 4. Magnetic flux density distribution of LS-SMPMSG

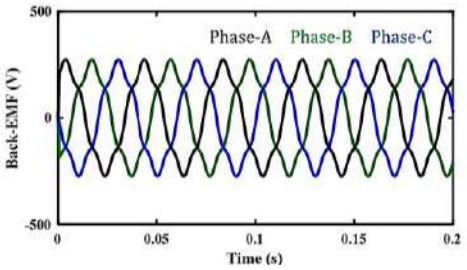


Figure 5. The LS-SMPMSG Back-EMF

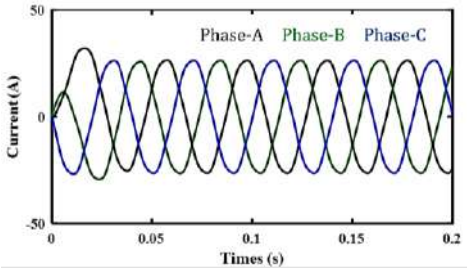


Figure 6. Three-phase stator currents in LS-SMPMSG

3. PROPOSED STRUCTURE FOR DESIGNED LS-SMPMSG

3. 1. Slot Configurations in LS-SMPMSG In this subsection, a novel slot structure for LS-SMPMSG has been proposed. The proposed scheme is facilitated in terms of LS-SMPMSG manufacture and assembly, notably its teeth and slots. In the proposed slot structure, the tooth shape has been modified and the interior space consists of cables with XLPE insulation and air gap. In the novel structure for the LS-SMPMSG, an attempt has been made to improve the previous structure. The novel proposed LS-SMPMSG slot structure is illustrated in Figure 7. As highlighted, the slot height is *W*, and five opening slots W_n are displayed. Distinct components, inclusive of the cable, XLPE insulation, air insulation, stator yoke, slot opening, air gap, and the rotor. In the last step, these five opening slot sizes are compared jointly to choose the finest opening slot. The five postures are tabulated in Table 5.

3. 2. Comparison of Five Slots Configuration In conformity with Figure 7 and Table 5, the endeavour

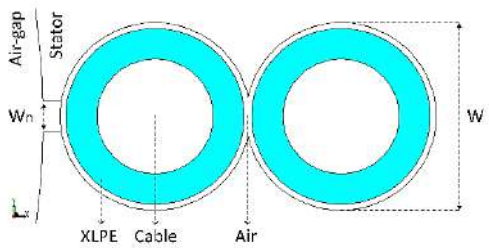


Figure 7. The proposed slot configuration for LS-SMPMSG

TABLE 5. Slot and opening slot dimensions

W_n (mm)	W (mm)	Ratio W_n / W
1	7	1/7
2	7	2/7
3	7	3/7
4	7	4/7
5	7	5/7

is to designate the incomparable opening slot for the new proposed slot configuration of the designed LS-SMPMSG. Many restrictions, such as maximum magnetic flux density, current, flux leakage, losses ordinarily Joule losses, torque, and coil inductance have been investigated to select the distinguished $W_n=1^{mm}$. Figure 8 illustrates the maximum magnetic flux density in the proposed slot configuration based on the opening slot of Table 8. In the first case ($W_n=1^{mm}$), the maximum flux density is achieved and the second state ($W_n=2^{mm}$) is the minimum. Forasmuch as the generator stator is wound, and cable is used as winding, the current passing through the stator is very indispensable. In this case, due to the shown results of Table 6, the maximum current has been obtained when the opening slot is equal to 5^{mm} which is equal to 27.17 A, and the lowest current is obtained in $W_n=1^{mm}$. The cable winding is precisely insulated with XLPE. This is one of the significant design points that should be considered in the LS-SMPMSG designing process.

The advantages of cable winding are high current and voltage tolerance. Due to the minuscule span and cross-sectional area inside the slot, the cables should be passed with great care to avoid damages to the LS-SMPMSG body and cable. Unquestionably, using the instruments for winding is the finest choice. The upper part of the slot has the same material as the LS-SMPMSG stator. This is deliberately chosen and has a positive effect on the performance of the LS-SMPMSG. In the proposed structure, in addition to having a smaller air gap, less iron is consumed. This factor has a direct impact on the weight of the machine, its manufacturing cost and its assembly cost. Joule losses, otherwise called

TABLE 6. Stator currents, Joule losses, torque and coil inductance in 5 different modes of the slot opening (W_n)

Slot Opening (mm)	Stator Currents (A)	Joule Loss (W)	Torque (Nm)	Coil Inductance (mH)
1	19.2	585.8	188.860	0.34
2	22	721.58	239.636	3.2
3	24.5	850.33	262.191	0.19
4	24.81	981	311.063	0.89
5	27.7	1071.41	343.559	4.7

RI^2 losses, are amidst the outstanding losses in the machines. One of the principal targets of such a design is to reduce losses in the machine. In agreement with Table 6, the lowest Joule losses occur in $W_n=1^{mm}$ with approximately 585.8 W. In the PM machines, the instantaneous electromagnetic torque depends on the air gap flux density waveform [16]. According to the increase in the slot opening, the torque is also increased, concerning Table 6. The findings of Table 6 exhibit that the LS-SMPMSG with $W_n=1^{mm}$ has a 188.86 *Nm* torque and increases until it reaches 343.559 *Nm* for $W_n=5^{mm}$. One of the implications which can be concluded is that a slight change in W_n leads to large changes in torque of LS-SMPMSG. Following Table 6, above all $W_n=1^{mm}$, the coil inductance is equal to 0.34×10^{-3} *H* which varies in different cases. For instance, 3.2×10^{-3} *H*, 0.19×10^{-3} *H*, 0.89×10^{-3} *H*, and 4.7×10^{-3} *H* for supplementary other four modes, respectively. To observe the changes in inductance, one of the coils is used and the diverse phases, experience the selfsame changes in the identical range of W_n alterations, according to Table 6 and also it has been found out that uneven variations of the coil inductance occur with increasing W_n. With regards to Figure 8, the maximum magnetic flux density appears in $W_n=1^{mm}$, and the minimum appears in $W_n=3^{mm}$. The high magnetic flux density gives rise to tooth saturation and it is a nuisance in the machine. According to Table 6, the stator embraces the minimum current in $W_n=1^{mm}$ and the maximum ones in the $W_n=5^{mm}$. An upward trend occurs with increasing W_n in the stator current. As it has been shown in Figure 9, the flux-linkage decreases with increasing W_n. Ultimately, the flux-linkage can be reviewing the latest significance. The $W_n=2^{mm}$ state is reduced 11.9% in comparison to $W_n=1^{mm}$, while this value between $W_n=3^{mm}$ and $W_n=2^{mm}$ is 13.5%. The comparison between $W_n=4^{mm}$ and $W_n=3^{mm}$ shows that $W_n=4^{mm}$ is 0.904 of the $W_n=3^{mm}$ and $W_n=5^{mm}$ is 0.922 of the $W_n=4^{mm}$. At last, $W_n=5^{mm}$ in comparison to $W_n=1^{mm}$ is decreased by 38.1%. In the small openings slot range, the leakage flux increases through the tooth tips which led to decreasing in flux linkage. By increasing the opening slot (W_n) at the range of greater than 4^{mm}, the flux linkage, and so, the internal machine voltage decreases. If the opening slot (W_n), not be in the proper ratio (too big or too small), the flux linkage decreases and the machine does not have its maximum rated voltage. Subsequently, like the assorted proposed slot modes, a proper opening slot should be selected as an LS-SMPMSG scheme which would be acceptable in all aspects. The opening slot (W_n) size is determined by the designer by using practical methods when manufacturing the machine. According to [17], the proper size of the opening slot for real machines is between 2.5–4^{mm} that involves numerous parameters such as the electromagnetic and the mechanical considerations of the machine. In this case, this parameter varies between 1-5^{mm} which is mostly about simulating and comparing results to determine how this parameter affects the performance of the LS-SMPMSG. This parameter has affected many parameters of the LS-SMPMSG. According to the simulation results, the proper dimension for the opening slot (W_n) is 3^{mm}. This value is a better choice that ensures the electromagnetic performance of the LS-SMPMSG and, based on this value, the improved design of the LS-SMPMSG is conducted.

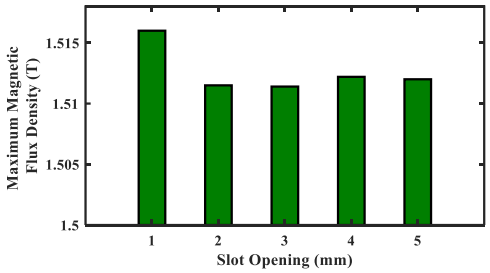


Figure 8. Distribution of maximum magnetic flux density in five different modes of the slot opening

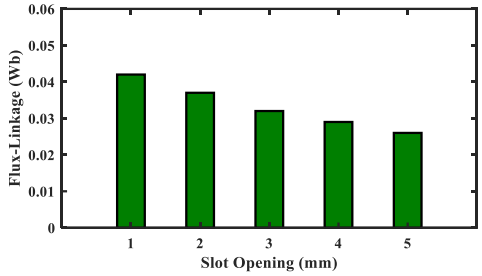


Figure 9. Flux linkage in five different modes of the slot opening

4. FEM INVESTIGATION OF THE NEW SLOT STRUCTURE

Figure 10 shows the 3D-view of the designed LS-SMPMSG with a new slot structure. In this part, the FEM analysis of the designed LS-SMPMSG is done. Figure 11 illustrates the flux line distribution of LS-SMPMSG with a new slot structure in the LS-SMPMSG stator. The stator flux density in the slot is reduced in *'New structure'*. As shown in Figure 12, the stator flux density is 1.10 *T* though in *'Old structure'* it was equal to 1.50 *T*.

5. COMPARISON BETWEEN THE NEW AND THE OLD SLOT STRUCTURE OF LS-SMPMSG

After comparing, analyzing, and examining the different structures for the opening slot, this section examines the

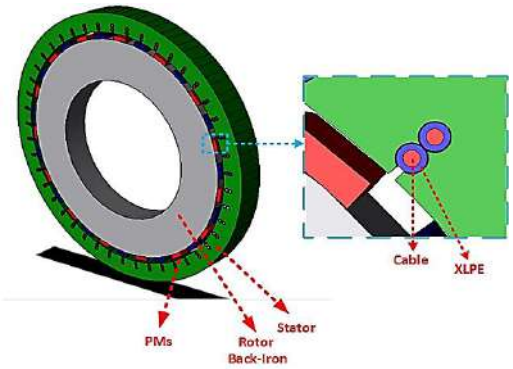


Figure 10. 3D-view of new designed LS-SMPMSG with a new slot structure

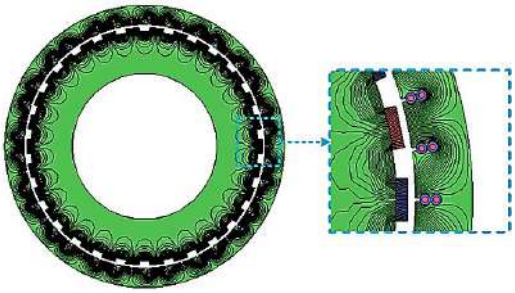


Figure 11. Magnetix flux line of new designed LS-SMPMSG with a new slot structure

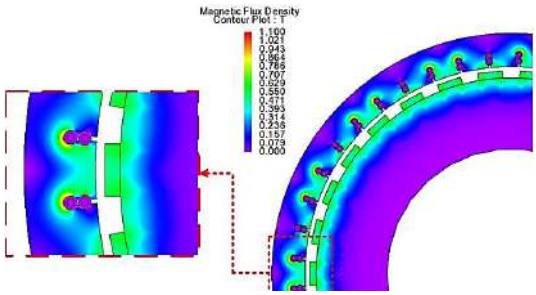


Figure 12. Magnetic flux density distribution of new designed LS-SMPMSG with a new slot structure

parameters of the new slot configuration of LS-SMPMSG with the previous ones. The previous structure with a circular slot configuration is named *'Old structure'* and the new structure with a modified slot configuration is labelled as *'New structure'*. Figure 13 shows the LS-SMPMSG torque in two forms, in the *'New structure'* the torque ripple is much less than '*Old structure'*. One more thing is the cogging torque, which depicts a significant change in the *'New structure'*. Cogging torque is a result of the interaction of PM and slots. Cogging torque, despite the torque ripples, results in noise and vibration when the machine is under load. The acoustic noise resulted from torque ripple is deleterious for the system in all cases. Changes in cogging torque are understandable in Figure 14. As it is observable, the cogging torque of the *'New structure'* has been compared to the *'Old structure'*. It is noteworthy that cogging torque has been drastically reduced by 87.5% which is a pleasant development for the *'New structure'* of LS-SMPMSG.

An extraordinarily feature of the proposed LS-SMPMSG in this paper is its slight cogging torque. In the new LS- SMPMSG structure, the design of the stator has been altered so that the redundant parts of the stator body which acts as an external air gap in the machine are eliminated. This design led to a reduction in the THD and the ripple of LS-SMPMSG torque. The significant losses are due to the Joule losses in the armature. In the *'New structure'* of LS-SMPMSG, the focus is on reducing losses, notably the Joule losses which have been demonstrated in Figure 15. In the *'New structure'* of LS-SMPMSG, the Joule losses have been

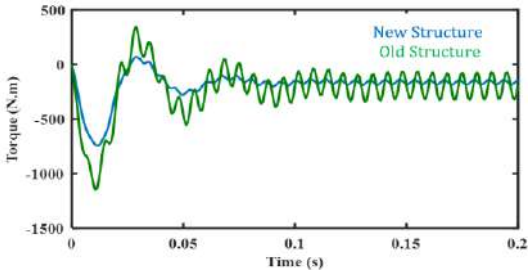


Figure 13. Torque in the *'Old structure'* and the *'New structure'* of LS-SMPMSG

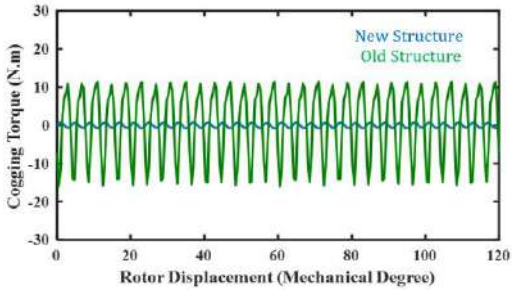


Figure 14. Cogging torque in the '*Old structure'* and *the 'New structure'* of LS-SMPMSG

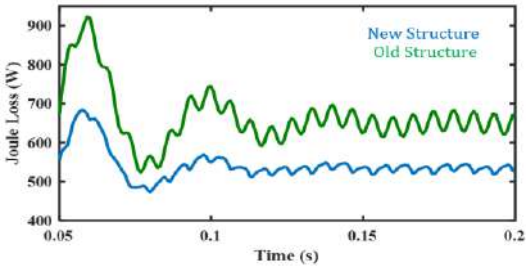


Figure 15. Joule loss in the *'Old structure'* and the *'New structure'* of LS-SMPMSG

TABLE 10. Weight in the *'Old structure'* and *the 'New structure'* of LS-SMPMSG

Old slot structure	New slot structure
292.03 Kg	291.60 Kg

diminished by 16.1%. To examine the changes made in the LS-SMPMSG, the weight of the machine has been examined for two structures. Table 10 compares the entire weight of LS-SMPMSG, proportionately. The results show that the *'New structure'* has reduced the weight of the LS-SMPMSG by 0.147%.

6. CONCLUSION

In this paper, a new slot structure has been presented based on five alternating opening slots in the LS-SMPMSG stator. The torque ripple and the cogging torque have been reduced by 19.54% and 87.5% respectively. In addition, the Joule losses and the weight of the LS-SMPMSG have been reduced by 16.1% and 0.147%, respectively. The outcomes show that the novel LS-SMPMSG slot structure proposed in this paper is very accomplishable in practice.

7. REFERENCES

1. He, C., Wu T., "Analysis and design of surface permanent magnet synchronous motor and generator,", ***CES Transactions on Electrical Machines and Systems***, Vol. 3, No. 1, (2019), 94-100. DOI: 10.30941/CESTEMS.2019.00013
2. Sun, X., Hu, C., Zhu, J., Wang, S., Zhou, W., Yang, Z., Lei, G., Li, K., Zhu, B., Guo, U., ''MPTC for PMSMs of EVs with multi-motor driven system considering optimal energy allocation,'' ***IEEE Transaction of Magnetic***, Vol. 55, No. 7, (2019), Art. no. 8104306, DOI: 10.1109/TMAG.2019.2904289
3. Shimoda, T., "A prospective observation of bonded rare-earth magnets.", ***IEEE Translation Journal on Magnetics in Japan***, Vol. 8, No. 10, (1993), 701-710. DOI: 10.1109/TJMJ.1993.4565726
4. Lin, Q., Niu, S., Fu, W. N., "Design and Optimization of a Dual-Permanent-Magnet Vernier Machine With a Novel Optimization Model.", ***IEEE Transactions on Magnetics***, Vol. 56, No. 3, (2020), DOI: 10.1109/TMAG.2019.2956071.
5. Pan, W., Li, W., Cui, L. Y., Li, X. M., Guo, Z. H., "Rare earth magnets resisting eddy currents." ***IEEE Transaction of Magnetics,*** Vol. 35, No. 5, (1999), 3343-3345. DOI: 10.1109/20.800519
6. Rahman, M. A., "History of interior permanent magnet motors [history]," ***IEEE Industry Applications Magazine***, Vol. 19, No. 1, (2013), 10–15. DOI: 10.1109/MIAS.2012.2221996
7. Patel, A.N, Suthar, B.N., "Design optimization of axial flux surface-mounted permanent magnet brushless dc motor for electrical vehicle based on genetic algorithm", ***International Journal of Engineering, Transactions A: Basics,*** Vol. 31, No. 7, (2018), 1050-1056. DOI: 10.5829/ije.2018.31.07a.07
8. Culina, J., Karsten, R., "Comparison of different resolution models and observed current profiles in the Bay of Fundy, Canada using turbine-relevant metrics," 9th European Wave and Tidal Energy Conference., Southampton, (2011). DOI: 10.1007/s00773-018-0601-z
9. O'Sullivan, D. L., Lewis, A., "Generator requirement and Functionality for Ocean Energy Converters," The XIX International Conference on Electrical Machines - ICEM 2010**.,** 1-7., (2010). DOI: 10.1109/ICELMACH.2010.5608132
10. Arish, N., Teymouri, V., "Development of Linear Vernier Hybrid Permanent MagnetMachine for Wave Energy Converter," ***International Journal of Engineering, Transactions B: Applications,*** Vol. 33, No. 5, (2020), 805-813. DOI: 10.5829/ije. 2020.33.05b.12
11. Benelghali, S., Benbouzid, M. E. H., Charpentier, J. F., "Marine tidal current electric power generation technology: State of the art and current status," 2007 IEEE International Electric Machines & Drives Conference,. (2007), DOI: 10.1109/IEMDC.2007.383635
12. Aly, H. H. H., "A novel approach for harmonic tidal currents constitutions forecasting using hybrid intelligent models based on clustering methodologies," ***Renewable Energy***, Vol 147, Part 1, (2020), 1554-1564. DOI: 10.1016/j.renene.2019.09.107
13. Fekri, H., Shamsi-Nejad, M. A., Hasheminejad, S. M., "Performance Analysis of a Novel Three-phase Axial Flux Switching Permanent Magnet Generator with Overlapping Concentrated Winding", ***International Journal of Engineering, Transactions B: Applications***, Vol. 32, No. 2, (2019), 286-295. DOI: 10.5829/ije.2019.32.02b.14
14. Vagati, A., Pellegrino, G., Guglielmi, P., "Comparison between SPM and IPM motor drives for EV application," The XIX International Conference on Electrical Machines- ICEM 2010 (2010), 1-6, DOI: 10.1109/ICELMACH.2010.5607911
15. Thomas K, Grabbe M, Yuen K, Leijon M. "A low-speed generator for energy conversion from marine currents—experimental validation of simulations" ***Proceedings of the Institution of Mechanical Engineers, Part A: Journal of Power and Energy,*** Vol. 222, No. 4, (2008), 381-388. DOI: 10.1243/09576509JPE567
16. Jahns, T. M., "Torque production in permanent-magnet synchronous motor drive with rectangular current excitation," ***IEEE Transactions on Industry Applications,*** Vol. IA-20, No. 6, (1984), DOI: 10.1109/TIA.1984.4504490
17. Fu, F. and Tang, X., *"Induction machine design handbook,"* China Machine Press, 2002, ISBN: 7-111-09078-0.

Persian Abstract

چکیده

در این مقاله به طراحی الکترومغناطیسی ژنراتور سنکرون مغناطیس دائم آهنربای سطحی سرعت پایین (LS-SMPMSG) با استفاده از روش مبتنی بر المان محدود (FEM) پرداخته شده است. ژنراتور طراحی‌شده برای استفاده در سیستم‌های تبدیل انرژی دریایی (MEC) و سیستم‌های تبدیل انرژی جزومدی (TEC) مناسب است. ویژگی اصلی LS-SMPMSG پیشنهادی، قابلیت تولید توان الکتریکی با سرعت کم در حدود ۱۰۰ دور در دقیقه است. از جمله ویژگی خاص دیگر این ساختار، استفاده از کابل به جای سیم پیچ در استاتور LS-SMPMSG است. در بخش بعدی این مقاله، یک ساختار جدید برای شیارهای استاتور LS-SMPMSG به منظور ارتقا عملکرد آن ارائه شده است. نتایج بدست آمده از روش المان محدود نشان می‌دهد که ساختار جدید منجر به کاهش ریپل گشتاور به مقدار ٪۱۹.۵۴ و گشتاور دندانه ای به مقدار ٪۸۷.۵ درصد نسبت به ساختار قبلی LS-SMPMSG شده است. این ساختار جدید همچنین بر روی تلفات ژول در LS-SMPMSG تاثیر گذاشته است به نحوی که در طراحی جدید، تلفات ژول LS-SMPMSG نسبت به حالت قبلی ٪۱۶.۱ کاهش پیدا کرده است. این پیکربندی شیار از نظر ساخت و مونتاژ شرایط را برای طراحی LS-SMPMSG مناسب‌تر کرده است. نتایج بدست آمده از روش المان محدود مزیّت های ساختار جدید پیشنهادی برای LS-SMPMSG را تایید می کند.

International Journal of Engineering

Journal Homepage: www.ije.ir

Effect of Additional Zirconia on Fracture Mechanics of Bioactive Glass-ceramics using Digital Image Correlation

I. K. Sabree, M. A. Aswad*, H. S. Abd Ali

Department of Ceramics Engineering and Building Materials, Faculty of Materials Engineering, University of Babylon, Babylon, Iraq

PAPER INFO

Paper history:
Received 23 April 2021
Received in revised form 23 May 2021
Accepted 21 June 2021

Keywords:
Bioglass-ceramic
Zirconia Nanoparticles
Composite Material
Fracture Toughness
Crack Opening Displacement
Digital Image Correlation

ABSTRACT

Bioactive glass-ceramic is used as a replacement material for bone tissue due to its compatibility, bioactivity, and the ability to form a crystallized hydroxyapatite layer, which is similar in composition and structure to the inorganic component of the bone mineral phase. In this paper, bioglass-ceramic was toughened using zirconia to improve its mechanical behaviour (i.e. crack opening dsplacement and fracture toughness). The fracture toughness of the bioactive glass cerami/zirconia composite was measured using three-point bending technique. Digital Image Correlation (DIC) technique was utilized for visualized the crack initiation and calculation the crack opening displacement (COD) at the tip and mouth of the crack and measuring of the crack propagation of the bioactive glass cerami/zirconia composite. The results indicated that the incorporation of the zirconia particles improved the measured fracture toughness of the BGC/ZrO_2 composite. The toughness of composite bioceramics is enhanced due to crack branching and crack deflection due to the presence of zirconia particles.

doi: *10.5829/ije.2021.34.09c.02*

NOMENCLATURE

COD	Crack Opening Displacement (mm)	SEM	Scanning Electron Microscope
BGC	Bioactive glass-ceramic	K_{IC}	Fracture Toughness (GPa.$m^{1/2}$)
XRD	X-ray Diffraction	CMOD	Crack Mouth Opening Displacement (mm)
ZrO_2	Zirconia	CTOD	Crack tip Opening Displacement (mm)

1. INTRODUCTION

Bioactive glass-ceramic is groups of osteoconductive biomaterial that assessed for restorating and repairing of body tissues especially for orthopedics and dental implant. According to the location and function of the damaged tissue, glass ceramics may require high strength and fracture toughness to be a suitable repairing material [1]. The central point limiting the use of bioactive ceramics is their low mechanical strength and fracture toughness. One approach to use ceramics as implants in load-bearing applications is reinforcing the ceramic with a second phase [2]. The mechanical properties of bioglass are controlled by the addition of oxides such as magnesia, alumina, zirconia, or titania. The use of composites of ZrO_2 and bioactive glass in the Na_2O-CaO-SiO_2-P_2O_5 system [3], where high strength comes from the zirconia reinforcing process and good compatibility and high bioactivity come from the bioactive glass, is an appealing way of producing strong bioactive materials [4]. Gali et al. [5] gave an evidence about increasing hardness and toughness of the glass ceramic-YSZ composites with varying amounts of YSZ (0, 5, 10, 15 and 20 wt.%). Kasuga et al. [1] confirmed the high bending strength of zirconia-toughened glass-ceramic composite with no degraded after in vivo implintation for 12 weeks, the optimum zirconia content to get high-strength and bioactivity was 30 vol.%. Rabiee and Azizian [6] used composite coating layer of bioactive glass-ceramic with various zirconia concentrations; the hardness test demonstrate that increasing zirconia content lead to rising of coating hardness. Zirconia was used to resolve the issue of ceramic brittleness and the resultant potential failed implants [7]. Fracture mechanics is a subject of

*Corresponding Author Email: mohsin.aswad@gmail.com (M. A. Aswad)

Please cite this article as: I. K. Sabree, M. A. Aswad, H. S. Abd Ali, Effect of Additional Zirconia on Fracture Mechanics of Bioactive Glass-ceramics using Digital Image Correlation, International Journal of Engineering, Transactions C: Aspects Vol. 34, No. 9, (2021) 2053-2059

engineering science that deals with the failure of solids caused by crack initiation and propagation [8], cracks are everywhere around us, these cracks which often exist could result in from industrial flaws or a variety of environmental conditions in the course of loading. Crack opening displacement and crack propagation are the most important parameters in fracture mechanics [9]. Fracture mechanics is depended on the stress intensity factor, wich is describes the stress concentration at the crack tip [10]. Measuring the crack opening displacement is difficult and as values obtain smaller, distinct equipment is needed and for a ceramic, SEM has been utilized [11]. These methods require precise sample preparations and distinctive care for determining the precise value of displacements [12-13]. To solve these challenges, non-contact optical technologies such as (DIC) Digital Image Correlation, which are used to determine the displacements and strains of samples for detecting crack propagation and measuring crack opening displacement, can be used to acquire a procedure for measuring the displacement of a specific location and precise strain measurements [14-15]. The aim of the study to improve the mechanical behaviour (i.e crack opening displacement and frcature toughness) of the bioactive glass ceamic tougheneing using zirconia and using digital image correlation for visualization of the crack initiation and measuring crack opening displacement and propagation in BGC/ ZrO_2 composite.

2. EXPERIMENTAL PROCEDURES

2. 1. Materials Preparation Bioglass-ceramic was prepared by melting technique. The preparation method and characterization were explained by Aswad et al. [16] and zirconia-yttria nanopowder/ nanoparticles (ZrO_2-3Y, 99.95%, 20 nm, metal basis).

2. 2. Bioglass-ceramic/ Zirconia Composite The composites were prepared from mixing different amounts of zirconia-yttria nanopowder/ nanoparticles (ZrO_2-3Y, 99.95%, 20 nm, metal basis) with prepared bioglass-ceramic powders. Where was mixed (3,5) wt.% of ZrO_2 nanopowder with bioglass-ceramic powder by using magnetic stirrer for 6 h using ethanol as solvent. The resulting slurry dried at 120 ^oC in the oven for 48 h. The dried powder was then crushed and sieved to be ready for compaction.

2. 3. Compact of BGC ansd BGC/ZrO_2 Specimens Bioglass-ceramic and BGC/ ZrO_2 powders were mixed with (2 wt %) poly vinyl alcohol, PVA as a binder's material [17]. Uniaxial semi-dry pressing techniques were used to form green ceramics specimens utilizing rectangles molds made of stainless steel with (60×6×5 mm) dimension. Figure 1 shows the compacts specimen for the fracture test. The appropriate pressure was 150 MPa. Solid specimens were bonded by sintering at temperatures 1000 °C with the heating rate (10°C/min) and soaking time 3 h and cooling down inside the furnace.

2. 4. Single Edge Notched Beam Sample Figure 2 shows the sample with dimensions (3 ×4 ×50) mm, and the samples were prepared to study fracture using a single edge notch beam [18]. Notch was made using a 170 μm thick diamond cutting disk (BT-MS210) in the ceramic laboratory/ Material Engineering College/ Babylon University [19].

2. 5. Digital Image Correlation The DIC method is an optical technique, which utilizes the full field, non-contact and high precision measurement of deformations and displacements. The artificial speckles pattern was created by sprayed randomly with black paint (vinyl acetate/ethylene VAE) on the BGC/ ZrO_2 composite specimen surface as shown in Figure 3. The speckles pattern are required to be non-repetitive, isotropic and contrast enough to permit the software to be able to identify and match the image before and after deformation [20-21]. The sprayed of a sample is necessary to make a subset, as shown in Figure 4. The software of DIC used for calculation named GOM (Gesellschaft fur Optische Mebtechnik) from a German company. These calculations start with a reference image captured before loading that was compared with other images during the loading periods. The assumption

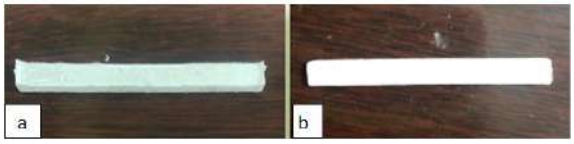


Figure 1. show compacts specimens for fracture test (a) before sintering, (b) after sintering with surface finishing (grinding and polishing)

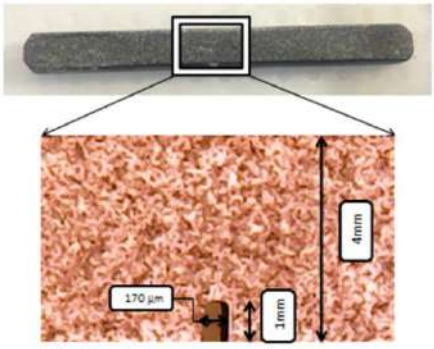


Figure 2. BGC/ ZrO_2 composite specimen after created Notched using SENB method

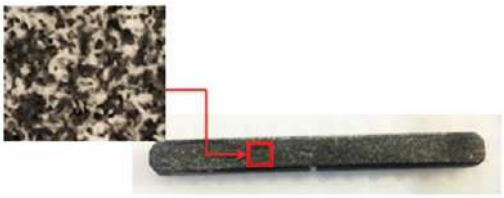


Figure 3. Speckles of black paint

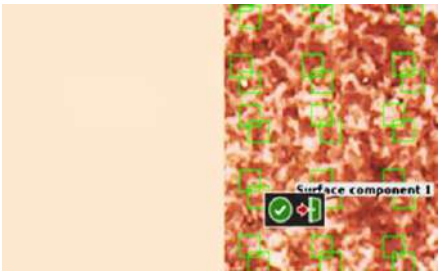


Figure 4. Subsets show when speckles of black paint (right side), unlike without black paint (left side)

proposed was that the color values of the image remain the same before and after the deformations. Firstly, the images are divided into subsets and look for the corresponding subset after deformations based on the assumptions and calculate their displacement; finally, deformations or displacement distributions maps are made. Figure 5 shows using digital image correlation for displacement measurement.

2. 6. Loading Setup and Test Procedure Specimens were subjected to the load microcomputers controlled electronics universal testing machines. These testing machines have loadings capacities of 5 KN and the loading rates are very low, resulting in a displacement not exceeding 0.005 mm/min. This camera is a digital microscope camera (Genesys Logic), and it has a lamp that is used to lighten the speckle pattern and is adapted to capture digital images during the loading process. The camera was positioned such that its lens was as parallel to the specimen surface as possible, and the focal length was changed to ensure that the picture was clear. The resolution of the camera was set to 640 × 480 pixels and the length pixel ratios of the imaging systems are 0.0008 mm/pixel. Cameras were programmed for capturing the image automatically at frames rates of 30 images/s, and these frames rates suit to store and to capture large numbers of images for further. The DIC method is to match maximum correlation between small zones (or subsets) of the specimen according to equations below:

$$R(x, y, x^*, y^*) = \sum F(x, y) - G(x^*, y^*) \quad (1)$$

$$C(x, y, x*, y*) = \frac{\sum F(x,y)\, G(x*,y*)}{\sqrt{\sum F(x,y)^2 \sum G(x*,y*)^2}} \quad (2)$$

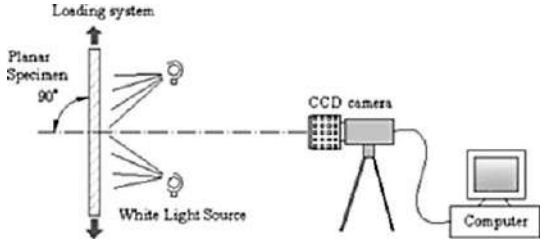


Figure 5. Setup for displacement measurement using digital image correlation [20]

2. 7 Measurement of The Crack Opening Displacement The position of crack mouth and crack tip opening displacement can be noticed in Figure 6 [22]. The crack opening displacement (CMOD and CTOD) can be measured throw DIC using two curves as shown in Figure 7. By the crack opening calculation with distance checks, you can analyze how the distance between two curves deviates from a reference value. The reference value can be a nominal value or a distance in the first active stage.

3. RESULTS AND DISCUSSION

3. 1. Fracture Toughness The fracture toughness of samples is known to be affected by the crack propagation behavior of composites [23]; thus, crack deflection and crack branching increase fracture toughness in the presence of ZrO_2 particles in the matrix as shown in Table 1. As the crack front interacts with a second-phase inclusion, such as ZrO_2 particles, the crack propagates out of the plane, lowering the stress intensity factor at the crack tip. To put it another way, it slows down crack propagation because it takes more energy to propagate a crack [24-25]. Crack deflection may be caused by residual stress in the composite or weak matrix/second phase interfaces [26].

Figure 6. Position of CMOD and CTOD in GOM

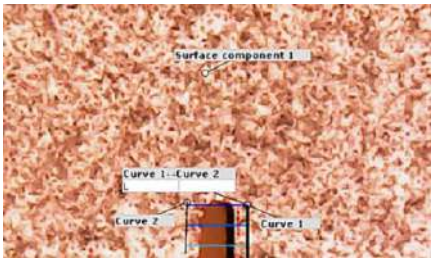


Figure 7. Measurement of CMOD and CTOD

TABLE 1. Fracture toughness of BGC and composites

Sample	Fracture Toughness (MPa.m$^{1/2}$)
Pure BGC	0.75
BGC / 3%ZrO_2	1.04
BGC / 5%ZrO_2	1.61

The fracture toughness of this study was calculated according to ASTM C1421 as shown below:

$$K_{Ic} = g\left[\frac{P_{max}\ S_o\ 10^{-6}}{BW^{3/2}}\right]\left[\frac{3[a/w]^{1/2}}{2[1-a/w]^{3/2}}\right] \quad (3)$$

3. 2. Toughening Mechanism The toughening mechanisms were investigated by Vickers' indentations were made on the polished surfaces and the cracks propagations path was as shown in Figure 8. The cracks of the BGC specimen show a transgranular fracture behavior, indicating that the toughness of BGC is independent of its grain size. Because of this transgranular fracture behavior, very straight cracks are often observed radiating from indentation corners, as shown in Figure 9, for the sample with 5wt.% ZrO_2. The crack is deflected when it encounters zirconia particles. Sometimes crack branching occurs by zirconia particles. Therefore, the toughening is attributed to both crack deflection and crack branching at sites of zirconia particles [25-27].

3. 3. Visualization and Propagation of Cracks DIC is a good method for visualized the crack and its initiation, otherwise, these cracks are not visible without using the DIC technique [28]. The results obtained from DIC to identified crack propagation with load and time of pure BGC and BGC-3, 5wt.% ZrO_2 is shown in Figure 10, respectively. These figures show the change in strain mapping near the notch at different loads and times for bioglass-ceramic and composite samples. Each figure

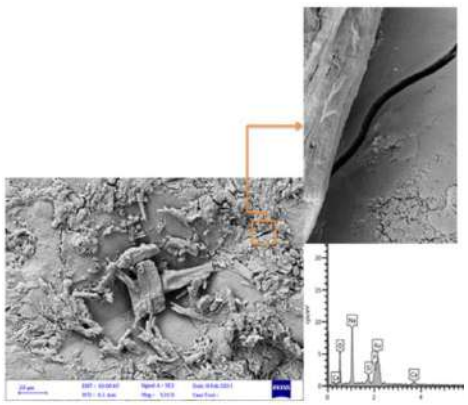


Figure 8. Indentation crack propagation pattern in pure bioglass-ceramic sample sintered at 1000°C

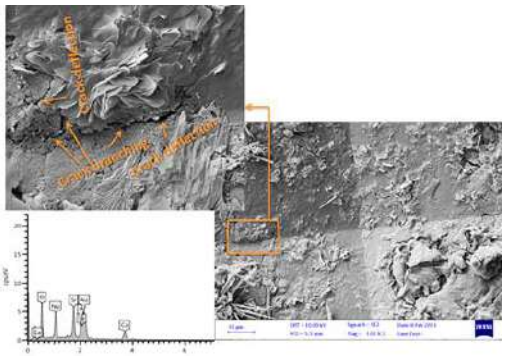


Figure 9. Indentation crack propagation pattern in BGC/5wt.% ZrO_2 sample sintered at 1000°C

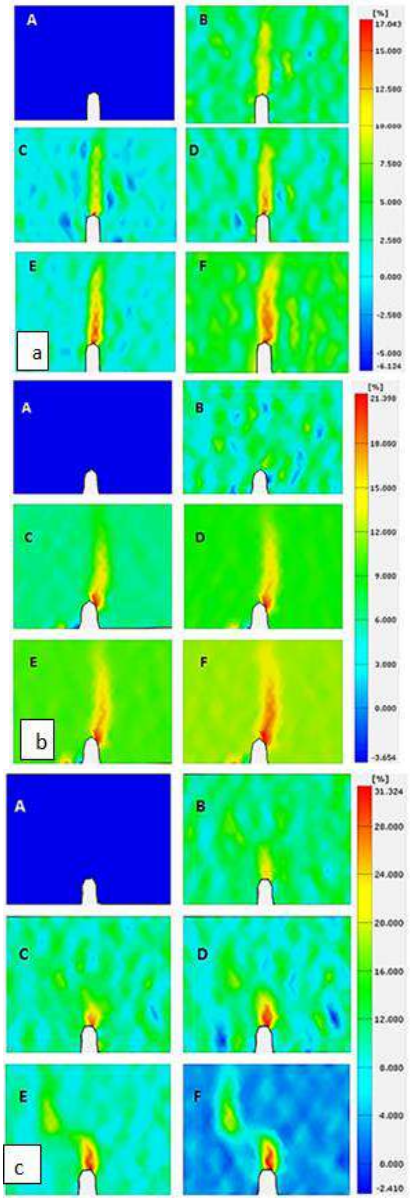


Figure 10. Strain maps with different load and time (a) Pure BGC,(b) BGC-3wt.% ZrO_2, (c) BGC-3wt.% ZrO_2

contains six images, The image (a) in every figure shows the undeformed sample (before applied loading), image (b) doesn't show a change in strain mapping near the notch tip, whereas the image (c) shows change in strain mapping near the notch tip which indicates to stress concentration near the notch tip. Critical load and crack have been formed as shown in image (d). Thereafter the specimen losing the bearing capacity and load rapidly decrease when the crack propagated as shown in images (e) and (f).The critical load for pure BGC (11 N) and the time when crack initiation was 412s. Also from this figure can be noticed that the critical load for (BGC-3wt.% ZrO_2) is 15 N at critical time 712s while the critical load for (BGC-5wt.% ZrO_2) is 22 N at critical time 1637s. The effect of addition ZrO_2 to bioglass-ceramic enhanced fracture toughness of bioglass-ceramic.

Figure 11 shows the relationship between load and time, in the SENB method. The curved divides into three parts: The first part represents the lower slope, it is before and the beginning of the emergence of the crack, The second part is steeper slope represents crack propagation before reaching the critical load, the top peak is represented a critical load. The last part is after the arrival of the sample to the critical load and continues to progress rapidly until the specimen fails.

3. 4. Crack Opening Displacement of Specimen

The relationship between applied force and crack opening displacement (CMOD and CTOD) [29] shown in Figure 12. From this figure can notice the force versus COD curve can be divided into three parts. In the first part, the relationship is linear, the second part start as soon as the curve deviates from linearity. The crack opening begins to widen at a faster rate, signaling the onset of material damage. The load continues to rise in this section until the peak load value is reached, at which point the material's full loading capacity is reached. As the crack mouth opening displacement continues to increase and the load begins to decrease after the peak load level, the last part begins [30].

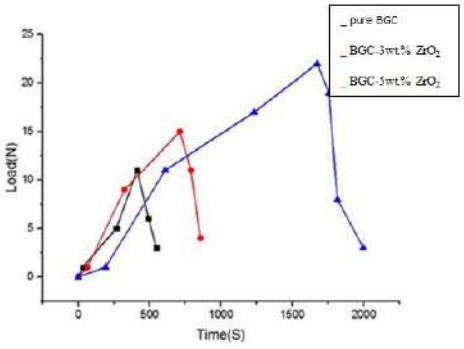


Figure 11. Represent the relationship between the load versus time of all specimens

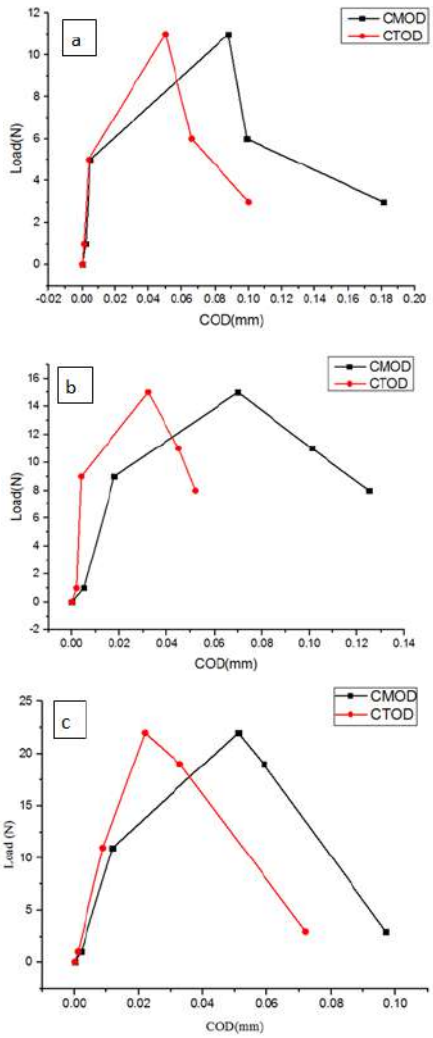


Figure 12. Load versus COD (a) BGC pure, (b) BGC-3wt.% ZrO_2, (c) BGC-5wt.% ZrO_2

4. CONCLUSIONS

According to the results of the experimental work, can be concluded the DIC is a good method for visualizetion of the crack initiation and measuring crack opening and propagation in BGC/ ZrO_2 composite. The critical stress intensity factor (K_{IC}) of BGC, BGC with 3, and 5wt.% ZrO_2 was 0.75, 1.04, and 1.61 MPa. $\sqrt{m}$ respectively, and the critical load of BGC, BGC/3, and 5 wt.% ZrO_2 were 11, 15, and 22 N, respectively for the experimental result.

5. REFERENCES

1. Kasuga. T., Yoshida. M., Ikushima A. J., M. Tuchiya, and Kusakar H.i, "Stability of zirconia-toughened bioactive glass-ceramics: in vivo study using dogs," ***Journal of Materials Science: Materials in Medicine***, Vol. 4, No. 1, (1993), 36-39, https://doi.org/10.1007/BF00122975.

2. Hench L. L., *An introduction to bioceramics*, Vol. 1. World scientific, (1993). https://doi.org/10.1142
3. Trang G. T., Linh N. H., and Kien P. H., "The study of dynamics heterogeneity in SiO_2 liquid," ***HighTech Innovation Journal,*** Vol. 1, No. 1, (2020), 1-7, DOI: 10.28991/HIJ-2020-01-01-01
4. Elgayar I., Aliev A. E., Boccaccini A. R., and Hill R. G., "Structural analysis of bioactive glasses," ***Journal of Non-Crystalline Solids***, Vol. 351, No. 2, (2005), 173-183, https://doi.org/10.1016/j.jnoncrysol.2004.07.06
5. Gali S., Ravikumar K., Murthy B. V. S., and Basu B., "Zirconia toughened mica glass ceramics for dental restorations," ***Dental Materials***, Vol. 34, No. 3, (2018), e36-e45, doi: 10.1016/j.dental.2018.01.009.
6. Rabiee S. M. and Azizian M., "Effect of zirconia concentration on the growth of nanowires in bioactive glass-ceramic coatings," ***International Journal of Applied Ceramic Technology,*** Vol. 10, No. 1, (2013), 33-39, https://doi.org/10.1111/j.1744-7402.2012.02844.x
7. Christel P. *et al.*, "Biomechanical compatibility and design of ceramic implants for orthopedic surgery," ***Annals of the New York Academy of Sciences***, Vol. 523, No. 1, (1988), 234-256, doi: 10.1111/j.1749-6632.1988.tb38516.x
8. Sun C. T. and Jin Z. H., "Fracture Mechanics," (2012), 7-8, doi: 10.1016/B978-0-12-385001-0.00001-8.
9. Nunes L. C. S. and Reis J. M. L., "Estimation of crack-tip-opening displacement and crack extension of glass fiber reinforced polymer mortars using digital image correlation method," ***Materials & Design***, Vol. 33, (2012), 248-253, https://doi.org/10.1016/j.matdes.2011.07.051
10. Dalmas D., Barthel E., and Vandembroucq D., "Crack front pinning by design in planar heterogeneous interfaces," ***Journal of the Mechanics and Physics of Solids***, Vol. 57, No. 3, (2009), 446-457, doi: 10.1016/j.jmps.2008.11.012.
11. Nguyen P. D., Dang V. H., and Eduardovich P. A., "Long-term Deflections of Hybrid GFRP/Steel Reinforced Concrete Beams under Sustained Loads," ***Civil Engineering Journal,*** Vol. 6, (2020), 1-11, DOI: 10.28991/cej-2020-sp(emce)-01
12. Lin Z. X., Xu Z. H., An Y. H., and Li X., "In situ observation of fracture behavior of canine cortical bone under bending," ***Materials Science and Engineering***, Vol. 62, (2016), 361-367, doi: 10.1016/j.msec.2016.01.061.
13. Mekky W. and Nicholson P. S., "The fracture toughness of Ni/Al2O3 laminates by digital image correlation I: experimental crack opening displacement and R-curves," ***Engineering Fracture Mechanics,*** Vol. 73, No. 5, (2006), 571-582.
14. Brynk T., Laptiev A., Tolochyn O., and Pakiela Z., "Digital Image Correlation Based Method of Crack Growth Rate and Fracture Toughness Measurements on Mini-Samples," ***Key Engineering Materials,*** Vol. 586, (2014), 96-99, doi: 10.4028/www.scientific.net/KEM.586.96.
15. Lyubutin P. S. and Panin S. V., "Mesoscale measurement of strains by analyzing optical images of the surface of loaded solids," ***Journal of Applied Mechanics and Technical Physics***, Vol. 47, No. 6, (2006), 905-910,. https://doi.org/10.1007/s10808-006-0131-z
16. Aswad M. A., Sabree I. K., and Abd Ali H S, "An examination of the effect of adding zirconia to bioactive glass-ceramic properties," *IOP* ***Conference. Series Material Science Engineering***, Vol. 1067, No. 1, (2021), 012121, doi: 10.1088/1757-899x/1067/1/012121.
17. Gharibshahian E., "The Effect of Polyvinyl Alcohol Concentration on the Growth Kinetics of KTiOPO4 Nanoparticles Synthesized by the Co-precipitation Method," ***HighTech Innovation Journal***, Vol. 1, No. 4, (2020), 187-193, Doi: 10.28991/HIJ-2020-01-04-06
18. Singh A., Kumar S., and Yadav H. L., "Experimental and Numerical Investigation of Fracture Parameters for Side Edge Notch Bend Specimen of Al 6063-T6," ***Iranian Journal of Science and Technology, Transactions of Mechanical Engineering,*** (2020), 1-19, https://doi.org/10.1007/s40997-020-00352-x
19. McCormack X., Wang J., Stover S. G. J., and Fyhrie D. "Effects of Mineral Content on the Fracture Properties of Equine Cortical Bone in Double Notched Beams," ***Bone,*** Vol. 50, No. 6, (2012), 1275-1280, doi: 10.1016/j.bone.2012.02.018
20. Pan B., Qian K., Xie H., and Asundi A., "Two-dimensional digital image correlation for in-plane displacement and strain measurement: a review," ***Measurement Science and Technology***, Vol. 20, No. 6, 062001, (2009), doi: 10.1088/0957-0233/20/6/062001.
21. Albo-Jasim M. M. H., Aswad M A., and Rashed H K., "Investigation of Crack Propagation and Opening in Hydroxyapatite Using Digital Image Correlation."***Journal of Engineering and Applied Sciences***, Vol. 12, (2017), 7953-7943, DOI: 10.3923/jeasci.2017.7935.7943 .
22. Panin S. V., Titkov V. V, and Lyubutin P. S., "Incremental approach to determination of image fragment displacements during vector field construction," ***Optoelectronics, Instrumentation and Data Processing,*** Vol. 50, No. 2, (2014), 139-147, https://doi.org/10.3103/S8756699014020058
23. Panin S. V., Maruschak P. O., Lyubutin P. S., Konovalenko I. V, and Ovechkin B. B., "Application of meso-and fracture mechanics to material affected by a network of thermal fatigue cracks," ***International Journal of Fatigue***, Vol. 76, (2015), 33-38, https://doi.org/10.1016/j.ijfatigue.2014.10.013
24. Faber K. T. and Evans A. G., "Crack deflection processes—I. Theory," ***Acta Metallurgica***, Vol. 31, No. 4, (1983), 565-576, https://doi.org/10.1016/0001-6160(83)90046-9
25. Rejab N. A., A. Azhar Z. A., Ratnam M. M., and Ahmad Z. A., "The relationship between microstructure and fracture toughness of zirconia toughened alumina (ZTA) added with MgO and CeO2," ***International Journal of Refractory Metals and Hard Materials***, Vol. 41, (2013), 522-530, https://doi.org/10.1016/j.ijrmhm.2013.07.002
26. Bengisu M. and Inal O. T., "Whisker toughening of ceramics: toughening mechanisms, fabrication, and composite properties," ***Annual Review of Materials Science,*** Vol. 24, No. 1, 83-124, 1994. https://www.annualreviews.org›
27. Abbas S., Maleksaeedi S., Kolos E., and Ruys A. J., "Processing and properties of zirconia-toughened alumina prepared by gelcasting," ***Materials*** *(Basel).*, Vol. 8, No. 7, (2015), 4344-4362, doi: 10.3390/ma8074344
28. Panin S. V, Chemezov V. O., Lyubutin P. S., and Titkov V. V, "Algorithm of fatigue crack detection and determination of its tip position in optical images," ***Optoelectronics, Instrumentation and Data Processing***, Vol. 53, No. 3, (2017), 237-244, https://doi.org/10.3103/S8756699017030062
29. Kaaim A. H.. Aswad M A., Awad S H, "Study on Iraqi bauxite ceramic reinforced Aluminum metal matrix composite synthesized by stir casting." ***International Journal of Engineering, Transactions A: Basics,*** Vol. 33, No. 7, (2020), 1331-1339. 10.5829/IJE.2020.33.07A.20
30. Marques B., Neto D., Antunes F., Vasco-Olmo J., and Díaz F., "Numerical tool for the analysis of CTOD curves obtained by DIC or FEM," ***Fatigue & Fracture of Engineering Materials & Structures,*** Vol. 43, No. 12, (2020), 2984-2997, https://doi.org/10.1111/ffe.13350

Persian Abstract

چکیده

سرامیک شیشه ای زیست فعال به دلیل سازگاری ، فعالیت زیستی و توانایی تشکیل یک لایه هیدروکسی آپاتیت متبلور ، که از نظر ترکیب و ساختار مشابه با اجزای غیر آلی فاز مواد معدنی استخوان است ، به عنوان ماده جایگزینی برای بافت استخوان استفاده می شود. در این مقاله ، سرامیک بیوگلاس با استفاده از زیرکونیا برای بهبود رفتار مکانیکی آن (به عنوان مثال محل انعطاف پذیری بازشو و مقاومت در برابر شکست) سخت شد. مقاومت در برابر شکست کامپوزیت شیشه زیستی فعال / سرامیک با استفاده از روش خمش سه نقطه ای اندازه گیری شد. از روش همبستگی تصویر دیجیتال DIC برای تجسم شروع ترک و محاسبه جابجایی بازشدگی ترک (COD) در نوک و دهان ترک و اندازه گیری انتشار ترک ترکیبی از سرامیک شیشه ای / زیرکونیا استفاده شده است. نتایج نشان داد که ترکیب ذرات زیرکونیا مقاومت شکستگی اندازه گیری شده کامپوزیت BGC / ZrO_2 را بهبود می بخشد. مقاومت بیوسرامیک کامپوزیت به دلیل انشعاب ترک و انحراف ترک به دلیل وجود ذرات زیرکونیا افزایش می یابد.

International Journal of Engineering

Journal Homepage: www.ije.ir

Graph Theoretic Loss Allocation Method for Microgrids having Variable Generation

D. Bharti*

Department of Electrical & Electronics Engineering, Ajay Kumar Garg Engineering College, Ghaziabad, UP, India

P A P E R I N F O

Paper history:
Received 8 August 2020
Received in revised form 09 April 2021
Accepted 07 July 2021

Keywords:
Loss Allocation
Microgrid
Variable Generation
Power Flow Solution
Relative Position

A B S T R A C T

With some assumptions and limitations, various methods have been developed in literature mainly for loss allocation in transmission network and afterwards extended for radial distribution network and some methods are specifically developed for radial distribution network. But, these methods are not suitable for microgrids which are integrated with conventional grid at sub-transmission and distribution levels depending on their geographical location. This paper presents a loss allocation method based on power flow results and relative position of buses for interconnected microgrid which is very effective in case of frequent change of generations due to intermittent nature of renewable resources. The implementation of the proposed method is very simple in microgrid with both meshed as well as radial topology without any computational complexity and requires only power flow solution and network data. The results are illustrated for different generating conditions of renewable sources in microgrid to demonstrate the efficiency of proposed loss allocation method.

***doi**: 10.5829/ije.2021.34.09c.03*

1. INTRODUCTION

Microgrid is an integration of various distributed generation (DG) especially renewable energy sources such as photovoltaic and wind which operates autonomously or in synchronous with conventional electrical grid. The incorporation of microgrids into distribution system has also transformed the structure of grid from radial to weakly meshed network. Energy insecurity, climate change and pollution are major concerns addressing significant changes in energy infrastructure by integrating renewable energy generation [1]. In modern power system structure, several renewable generations are integrated to conventional grid at sub-transmission level and several may be connected at distribution level. Hydro plant and wind farm are always far away from populated area and need to be connected to long-distance transmission [1]. Some of solar generations are present at low voltage distribution level. Due to independent ownership of DGs present in microgrid, it is essential to have a robust loss allocation (LA) method for attaining transparency. LA method should be applicable to both radial and meshed structured microgrid because depending on type of integration, some of the microgrids are of radial topology and some are of meshed topology.

1. 1. Literature Review In microgrids, conventional power generations are required to avoid power interruption as electricity generation by renewable energy resources are intermittent [1]; the output from a wind farm or a photovoltaic array depends on the climatic conditions. In interconnected mode, microgrid is connected with distribution network and works in co-ordination with the distribution management system. The power flow pattern varies very frequent in microgrids due to integration of non-conventional generation units. The presence of multiple source changes the distance between sources and loads which also alters network usages. Any LA method intended to be used for microgrid operation must be equally applicable to both radial as well as meshed network topology since the microgrid can be of either configuration. This makes LA problem very significant in microgrids. A robust LA method is required to differentiate between the contributions of individual participants i.e. the generators and loads connected in the

*Corresponding Author Institutional Email: bhartidibya@akgec.ac.in (D. Bharti)

Please cite this article as: D. Bharti, Graph Theoretic Loss Allocation Method for Microgrids having Variable Generation, International Journal of Engineering, Transactions C: Aspects Vol. 34, No. 9, (2021) 2060-2069

microgrid and it should also consider amount of network usage of any participant as well as distance from source while making allocations to it.

There exist various LA methods in literature for transmission systems [2]. With some assumptions and modifications, transmission LA methods can also be used for distribution systems. Table 1 summarized discussion about different allocation methods proposed earlier.

An improved average LA method for distribution network is proposed by Zhang et al. [3]; which is especially suitable for harmonic loss. Moret et al. [4] presented an analysis of LA policies for avoiding market outcomes which categorizes agents for geographical location. It also suggests inclusion of system operator in both transmission and distribtion level.

1. 2. Contribution of Proposed Work A fair and satisfactory LA method should reflect both the network topology and the magnitude of power injected or consumed at a bus. The present work proposes a LA strategy for interconnected microgrid with meshed as well as radial structure which works on power flow result of the system and relative distance between buses in network. Proposed method requires only power flow results and electrical closeness centrality indices, which is a measure of the degree to which an individual is near all other individuals in a network. The proposed LA method is straightforward without any intricate computational applications.

1. 3. Organization of Paper The next section confers electrical closeness centrality measures and their

TABLE 1. Different existing LA methods

Ref.	Method	Merits	Demerits
[2]	Pro-rata method	Allocation of loss is characterized by loss proportionally to the power delivered by generators and loads.	Neglects relative location of generators and loads within the network.
[5]	MW-mile methods	Considers the relative position of each participant from slack bus	Neglects the amount of power flowing through network.
[6]	ITL (incremental transmission loss) methods	Suitable for networks with high X/R ratio and dependent on choice of slack bus	Due to dependency on slack bus, ITL method results in over recovery of network loss.
[7]	DLC (direct loss coefficient) method	Allocates loss directly by establishing relation between real/reactive power of a bus and network loss	Application of Hessian included in procedure of DLC is computationally exhaustive for handling larger system
[8]	Proportional sharing principle based methods	Allocate total network loss to either generators/DGs or loads as it involves application of linear principle	Proportional sharing principle is an assumption.
[9]	Z-bus/Y-bus methods	Easy to implement in larger networks	Not applicable to microgrid with radial topology when shunt admittance of the lines are negligible
[10, 11]	Circuit theory based methods	Suitable for system with high value of X/R ratio	Not justified in microgrids connected at distribution level where X/R ratio is low
[12]	Branch current decomposition method	Suitable for radial distribution configuration with DGs	Requires an additional forward sweep power flow on modified network
[13]	Power summation method	Tracing based approach	Suitable for radial distribution system
[14]	Energy summation method	Based on disintegration of energy	Suitable for radial distribution system
[15]	Exact formulation method	Suitable for allocating branch loss to the nodes	Applicable to radial system
[16]	Branch oriented method	Loss are allocated to loads and DGs	Results into over-recovery of loss
[17]	Current/Power summation method	Easy to implement	Suitable for radial distribution system
[18]	Branch oriented methods	Employ backward sweep reduction technique	Applicable only for microgrid with radial topology
[19]	Game theory based methods	Overcomes the limitations of conventional Shapley value method	Considers DGs as negative loads
Proposed method	Pro-rata method	Allocates loss directly by establishing relation between real power of a bus and network loss. Considers relative position of each participants, Suitable for microgrid with radial & meshed topology, easy to implement	Applies normalization process for distributing loss to generators and loads.

use in proposed LA method. Section 2 discusses relevance of electrical closeness centrality measure in LA in brief and describes the steps of calculating closeness centrality and proposed method with an example. Section 3 demonstrates the application of proposed method in different scenario of microgrid and presents the comparison of proposed method with existing methods. The results of proposed method for larger microgrid are given in section 4. Finally, section 5 concludes the work.

2. CENTRALITY MEASURES AND LOSS ALLOCATION

Distributed energy resources (DERs) of microgrids are owned by different entity which necessitates implementation of a robust LA method by distribution system operator (DSO). In microgrids, generations are sporadic subject to climatic changes which lead to change in power flow results. With changing generation pattern the power flow through different paths of the network changes, resulting into change in network usage. Electrical closeness centrality is the measure of relative position of a bus in the network and dependent on system data and power flow results. To change the loss contribution according to relative position and network usage by individual participant, electrical closeness centrality is used for allocating loss to generators and loads.

Electrical closeness centrality measures are calculated by using bus dependency matrix [20], which exhibits dependability of buses on each other present in network. The method discussed by Bharti and De [20] for finding bus dependency matrix is applicable to both meshed and radial electrical network. A fair LA method needs to take care of the relative location and importance of any bus in the network and this aspect can be included by incorporating centrality index in the allocation method. The following section discusses the method of determining centrality measure of a network.

3. 1. Calculation of Electrical Closeness Centrality Measures

The electrical closeness centrality measure is calculated from the bus dependency matrix which is of order (bus*bus). The bus dependency matrix of any $'n'$-bus system is calculated as Equation (1) then:

$$D_{bus\_dep_n} = \begin{bmatrix} d_{11} & d_{12} & \dots & d_{1n} \\ d_{21} & d_{22} & \dots & d_{2n} \\ \vdots & \dots & \dots & \vdots \\ d_{n1} & d_{n2} & \cdots & d_{nn} \end{bmatrix} \quad (1)$$

The computation of bus dependency matrix depends on shortest path between pair of buses and active power flowing it. In shortest path between pair of bus of power system network, there will be intermediate buses if selected buses are not adjacent buses. There may be either single bus or multiple intermediate buses in shortest path. So, the elements of bus dependency matrix can be calculated as:

Step I: Run power flow of the system.
Step II: Determine the shortest path for each bus pair by assigning impedance as weights.
Step III: Find the maximum active power in each shortest path, P_{st}. (see Figures 1 and 2).
Step IV: Ascertain the maximum of inflow and outflow at intermediate bus within each shortest electrical path, $P_{st}(i)$. (see Figures 1 and 2).
Step V: The dependency of bus $'s'$ upon bus $'i'$ to transmit power to other buses of the network can be given by,

$$d_{si} = \sum^{n}_{\substack{t=1 \\ s\neq t\neq i\in V}} \frac{P_{st}(i)}{P_{st}}$$

For example, in a 5-bus system, element d_{14} and d_{35} are calculated as:

$$d_{14} = \sum_{t\in\{2,3,5\}} \frac{P_{1t}(4)}{P_{1t}} = \frac{P_{12}(4)}{P_{12}} + \frac{P_{13}(4)}{P_{13}} + \frac{P_{15}(4)}{P_{15}} \quad (2)$$

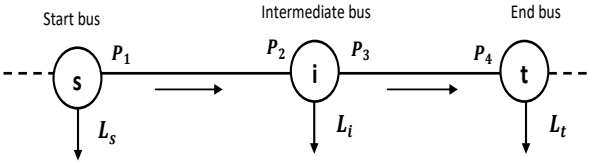
$$d_{35} = \sum_{t\in\{1,2,4\}} \frac{P_{3t}(5)}{P_{3t}} = \frac{P_{31}(5)}{P_{31}} + \frac{P_{32}(5)}{P_{32}} + \frac{P_{34}(5)}{P_{34}} \quad (3)$$

The diagonal elements of bus dependency matrix will be zero and its row summation gives electrical closeness centrality.

Electrical closeness centrality of each can be calculated by row sum of matrix. For example, closeness centrality of n^{th} bus will be calculated as:

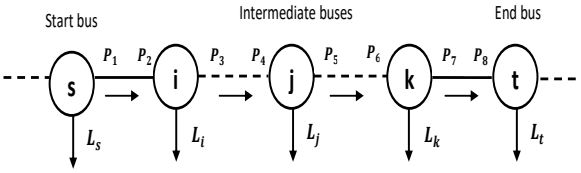
$$C_n = d_{n1} + d_{n2} + \cdots \dots \dots \dots + d_{nn} \quad (4)$$

The closeness centrality computes the extent of connectivity to which a bus is in close proximity to all other buses of the network. Electrical closeness centrality measures are calculated by using bus dependency matrix which is dependent on system impedance and power flow results. A fair LA method should incorporate the relative location and importance of any bus in the network and



$$P_{st} = \max of\ (P_1, P_2, P_3, P_4)$$
$$P_{st}(i) = \max of\ (P_2, P_3)$$

Figure 1. Description of $P_{st}(i)$ and P_{st} in shortest path with single intermediate bus



$$P_{st} = max\ of\ (P_1, P_2, P_3, P_4, P_5, P_6, P_7, P_8)$$
$$P_{st}(x) = max\ of\ [max(P_2, P_3), max(P_4, P_5), max(P_6, P_7)]$$

x is the intermediate bus depending upon amount of active power inflow/outflow.

Figure 2. Description of $P_{st}(i)$ and P_{st} in shortest path with multiple intermediate buses

this aspect can be included by integrating centrality index in the allocation method.

3. 2. Loss Allocation using Electrical Closeness Centrality Measures The algorithm used for allocating loss to each node of the network is given as follows:

Step I: Calculate total loss (P_{loss}) of the system using power flow results.

Step II: Calculate electrical closeness centrality indices (C_i) for each bus present in the network by using bus dependency matrix as discussed in section 2.1.

Step III: Calculate total outgoing power (P_i) at each bus present in the system.

Step IV: Calculate proportional indices (a_i) for each bus by using Equation (5).

$$a_i = \frac{P_i}{\sum_{i=1}^{n} P_i C_i} \tag{5}$$

Where, $'n'$ is the number of buses present in network.

Step V: Loss allocated to each bus can be given by

$$p_i = P_{loss}.a_i.C_i \tag{6}$$

Step VI: Normalize loss at various buses to calculate loss contribution of individual loads and generators.

The formula given as Equation (6) allocates loss to individual bus as:

$$p_i = P_{loss}.a_i.C_i$$

Here, a_i is proportional index of bus 'i' and C_i is electrical closeness centrality index of bus 'i'. p_i is loss allocated to bus 'i'. The proportional index of bus 'i', (a_i) depends upon outgoing power and closeness centrality of bus 'i' as represented in Equation (5). For calculating electrical closeness centrality by bus dependency matrix, shortest path between pair of buses and active power flowing in them will be determined (as detailed in section 2.1).

$$Where,\ a_i = \frac{P_i}{\sum_{i=1}^{n} P_i C_i}$$

And, $C_i = d_{i1} + d_{i2} + \cdots \ldots \ldots \ldots + d_{in}$

$$\Rightarrow C_i = \sum_{\substack{t=1 \\ i\neq t\neq 1\in V}}^{n} \frac{P_{it}(1)}{P_{it}} + \sum_{\substack{t=1 \\ i\neq t\neq 2\in V}}^{n} \frac{P_{it}(2)}{P_{it}} + .. + \sum_{\substack{t=1 \\ i\neq t\neq n\in V}}^{n} \frac{P_{it}(n)}{P_{it}}$$

So, allocation of loss to bus can also be written as:

$$p_i = P_{loss}.\left(\frac{P_i}{\sum_{i=1}^{n} P_i\left(\sum_{\substack{t=1 \\ i\neq t\neq 1\in V}}^{n} \frac{P_{it}(1)}{P_{it}} + \sum_{\substack{t=1 \\ i\neq t\neq 2\in V}}^{n} \frac{P_{it}(2)}{P_{it}} + \ldots + \sum_{\substack{t=1 \\ i\neq t\neq n\in V}}^{n} \frac{P_{it}(n)}{P_{it}}\right)}\right). \left(\sum_{\substack{t=1 \\ i\neq t\neq 1\in V}}^{n} \frac{P_{it}(1)}{P_{it}} + \sum_{\substack{t=1 \\ i\neq t\neq 2\in V}}^{n} \frac{P_{it}(2)}{P_{it}} + \cdots \ldots + \sum_{\substack{t=1 \\ i\neq t\neq n\in V}}^{n} \frac{P_{it}(n)}{P_{it}}\right) \tag{7}$$

In above expression,

P_{loss}: Total loss of the system using power flow results

P_i: Total outgoing power at each bus 'i' present in the system

n: Total number of bus in the system

P_{it}: Maximum active power in each shortest path

$P_{it}(k)$: Maximum of inflow and outflow at intermediate bus within each shortest electrical path

$P_{it}(k)$ will be 'zero' if shortest path is trough directly connected buses and 'non-zero' if there exists intermediate bus in shortest path.

3. 3. Validation of Proposed Loss Allocation Method A test system with 5-bus and 7 links is considered as microgrid with meshed topology to demonstrate the applicability of proposed method. Microgrid is integrated with conventional grid at bus 1 and there are two nonconventional energy resources: solar plant and wind farm with installed capacity of $40MW$ and $30MW$ respectively, connected at bus 2 and bus 3. Figure 3 represents 5-bus test system with system impedance, outflow power and inflow power at each node. The bus dependency matrix (given below in Equation (8)) is calculated which depends on shortest path and power flow result. Total loss of the system (P_{loss}) is $12.6806MW$.

$$D_{bus\_dep} = \begin{bmatrix} 0.0000 & 4.8650 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.9803 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 10.7646 & 0.0000 \\ 0.0000 & 1.9730 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 4.9730 & 0.0000 & 0.0000 & 0.0000 \end{bmatrix} \tag{8}$$

As described above, the row sum of bus dependency matrix can be used as electrical closeness centrality measures which are $C_1 = 4.8650$, $C_2 = 0.9803$, $C_3 = 10.7646$, $C_4 = 1.9730$ and $C_5 = 4.9730$ and outgoing power are $P_1 = 142.6805$, $P_2 = 149.3261$, $P_3 = 64.7077$, $P_4 = 111.1685$ and $P_5 = 60.0000$ for buses 1, 2, 3, 4 and 5 respectively.

By using Equation (6), loss allocated to buses 1, 2, 3, 4 and 5 can be calculated as $p_1 = 4.2837$, $p_2 = 0.9034$, $p_3 = 4.2985$, $p_4 = 1.3536$ and $p_5 = 1.8414$ respectively. The fairness of the LA method can be verified that $p_1 + p_2 + p_3 + p_4 + p_5 = 12.6806$ MW.

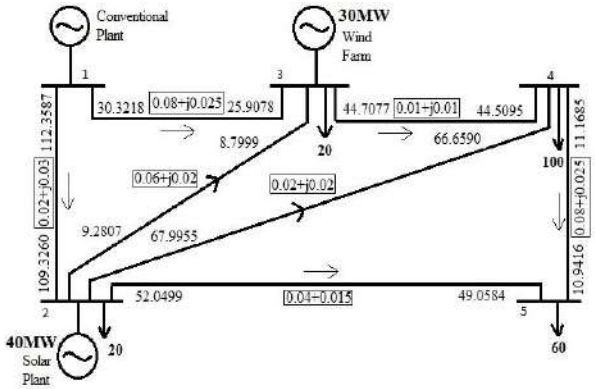


Figure 3. Meshed structured microgrid with 5 buses & 7 links

The proposed LA technique can be easily applied to the microgrids where change in generations is very frequent as it requires very less computational time and is a very fast and efficient technique. Electrical closeness centrality measures based LA method is efficiently applicable to both meshed structured microgrid and radial topology based microgrid.

4. APPLICATION OF PROPOSED METHOD

To illustrate applicability of proposed method with variable sources modified IEEE 14-bus test system is considered as an interconnected microgrid with meshed topology. Two different cases with altered renewable generation are considered to validate the applicability of proposed method in microgrids where alteration in power flow pattern is very frequent. The following scenarios are considered as microgrid for analysis of proposed LA method:

4. 1. Application of Proposed Loss Allocation Method With Variable Sources

Case 1: Interconnected microgrid with one solar plant and one wind farm

A modified IEEE 14-bus system is considered in which node 2 is assumed to be connected with a wind farm of 40MW rated capacity and node 3 has a concentrated solar plant of 60MW rated capacity. A new branch is added between buses 1 and 3 in this modified system and conventional grid is interconnected at bus1. Single line diagram of the modified test system is shown in Figure 4 with direction of power flow through the lines. Branch data for modified system are listed in Table 2.

After power flow, total loss of the system (P_{loss}) is calculated as 4.4700 MW. After power flow, by using LA technique proposed in section 2.2, loss allocated to each bus present in the network is given in Table 3.

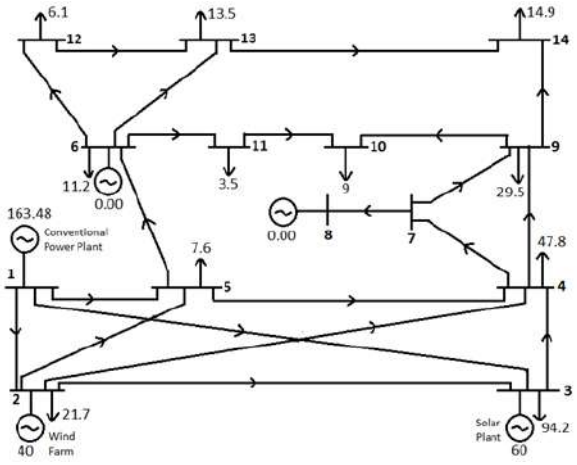


Figure 4. Modified IEEE 14-bus test system as meshed structured microgrid

TABLE 2. Branch data of modified IEEE 14-bus system

From Bus	To Bus	Resistance (in Ω)	Reactance (in Ω)
1	2	0.01938	0.05917
1	5	0.05403	0.22304
1	**3**	**0.00000**	**0.04211***
2	3	0.04699	0.19797
2	4	0.05811	0.17632
2	5	0.05695	0.17388
3	4	0.06701	0.17103
4	5	0.01335	0.04211
4	7	0.00000	0.20912
4	9	0.00000	0.55618
5	6	0.00000	0.25202
6	11	0.09498	0.19890
6	12	0.12291	0.25581
6	13	0.06615	0.13027
7	8	0.00000	0.17165
7	9	0.00000	0.11001
9	10	0.03181	0.08450
9	14	0.12711	0.27038
10	11	0.08205	0.19207
12	13	0.22092	0.19988
13	14	0.17093	0.34802

* Newly Added line

TABLE 3. Results of loss allocation of modified IEEE 14-bus system (with 2 DERs) considered as microgrid

Bus	Electrical Closeness Centrality	Outgoing Power (in MW)	Loss allocated (in MW)
1	9.9971	163.4800	1.0149
2	9.8958	96.4100	0.5924
3	10.8397	127.1300	0.8557
4	9.0000	94.3700	0.5274
5	8.9733	73.5900	0.4101
6	8.9777	41.6400	0.2321
7	9.9524	29.6100	0.1830
8	11.9524	0.0000	0.0000
9	9.9465	46.5700	0.2876
10	10.9456	9.0000	0.0612
11	10.9963	5.8200	0.0397
12	10.9993	7.5300	0.0514
13	9.9993	18.2100	0.1131
14	10.9407	14.9000	0.1012
	Total		4.4700

In Table 3, electrical closeness centrality, outgoing power and loss allocated to each bus is listed. In Table 3, loss allocated to bus 8 is zero as neither generator nor load is connected at bus 8 shown in (Figure 4). From Figure 4, it can also be observed that generator connected at bus 6 is not delivering any power but load is obtaining power. So, the loss is being allocated to bus 6. From results presented in Table 3, it can be concluded that total loss allocated to generators and loads are 2.463 MW and 2.0068 MW respectively. So, the proposed method is not dividing total loss to generators in equal proportional unlike pro-rata method and this is realized by including electrical closeness centrality indices which is representation of relative electrical distance of buses present in network.

Case II: Interconnected microgrid with two solar plants and two wind farms

Again, modified IEEE 14-bus system is considered as microgrid in which both node 2 and 6 are assumed to be connected with a wind farms of 40MW rated capacity and node 3 and 8 has a concentrated solar plant of 60MW and 50MW rated capacity respectively. After power flow, total loss of the system (P_{loss}) is calculated as 2.8777 MW. Proposed LA method is applied to IEEE-14 bus system considered as microgrid and results are listed in Table 4.

From results of Table 4, it can be concluded that total loss allocated to generators is 1.6671 MW while total loss allocated to loads is 1.2106 MW which indicates that total loss is not being shared by generators and loads in equal fraction which was drawback of LA methods related to pro-rata techniques. From Tables 3 and 4, it is clear that depending on power flow results, electrical closeness centrality changes. Loss allocated to buses changes depending on power flow results and electrical closeness centrality but sum of loss allocated to buses is exactly equal to the total loss of the system.

TABLE 4. Results of loss allocation of modified IEEE 14-bus system (with 4 DERs) considered as microgrid

Bus	Electrical Closeness Centrality	Outgoing Power (in MW)	Loss allocated (in MW)
1	10.8214	71.8779	0.3613
2	8.9686	86.0742	0.3586
3	10.9765	94.2000	0.4803
4	8.9801	58.9536	0.2459
5	8.9929	39.8391	0.1664
6	8.9440	45.3010	0.1882
7	9.9765	66.0341	0.3060
8	12.0000	50.0000	0.2787
9	9.9684	43.0276	0.1992
10	10.9715	9.0000	0.0459
11	10.9717	8.0126	0.0408
12	10.9748	7.7963	0.0397
13	9.9852	19.5840	0.0908
14	10.9666	14.9000	0.0759
Total			2.8777

4. 2. Application of Proposed Loss Allocation Method in Microgrid with Radial Topology A microgrid with radial topology is considered in Figure 5. It represents a modified 12-bus radial system integrated with conventional power plant at bus 1 and a solar plant and wind farm are connected at bus 5 and 9 respectively. It is assumed to be connected with a wind farm of 60MW rated capacity at bus 9 and node 5 has a concentrated solar plant of 90MW rated capacity. Bus 1 is not connected to any load; only conventional power plant is integrated into the network at bus 1. All the connected loads (shown in Figure 5) are in MW. Direction of power flow is shown in Figure 5.

After power flow, it has been found that the total loss of the system is 6.8429 MW. The loss allocated by the proposed method is listed in Table 5.

From results of Table 5, also it can be concluded that proposed method overcomes the drawback of LA methods based on pro-rata by sharing 2.9135 MW to generators and 3.9295 MW to loads.

It has been also observed from Tables 3, 4 and 5 that LA by proposed method has not given negative loss contribution to any generators or loads.

4. 3. Comparison of Proposed Loss Allocation Method For comparison of proposed method, two different test systems are considered: one with meshed and another with radial configurations. A 6-bus meshed topology with/without DGs connection can be found in literature [9] and 17-bus radial network with DGs can be perceived from [14].

4. 3. 1. 6-bus System with Meshed Topology A 6-bus test system with meshed topology, shown in Figure 6, is considered whose power flow results and system data can be found in literature [9]. Results obtained by proposed method for active power LA in 6-bus test

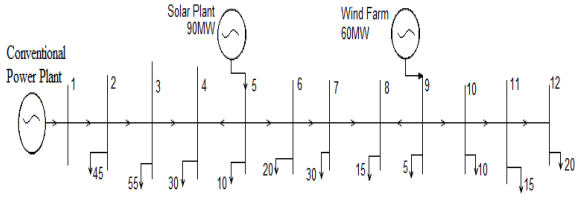


Figure 5. 12-bus radial system modified as microgrid

system is compared with the results of few other established and popular LA methods found in literature.

A modified 6-bus test system with meshed topology is considered where active power loss is 12.5561 MW. The proposed method for LA is applied to allocate loss contribution to each bus and then to loss contribution of individual generator and load. The comparison of LA to generators and loads with other methods for this 6-bus test system shown in Figure 6 is provided in Table 6.

From results listed in Table 6, it can be concluded that proposed method allocates 6.236 MW loss to generators connected at buses 1 and 2 while total loss allocated to loads is 6.3202 MW. In case of pro-rata method, total loss allocated to generators is 6.278 MW which is equal to total loss allocated to loads (6.278 MW). In cases of 6-bus system, method proposed in [9] gives negative loss contribution to bus 1 while proposed method does not allocate negative power loss contribution is any participants.

It can be observed from Table 6 that loss allocated to generators is 6.27811MW, 10.4771MW and 6.4231MW by pro-rata method [2], Z-bus method [7] and game theory method respectively. The proposed method allocates 6.236MW loss to generators, which seems similar as by pro-rata method [2]. Total loss allocated to loads are 6.278MW, 1.989MW and 6.106MW by pro-rata method [2], Z-bus method [7] and game theory method respectively. However, loss allocated to loads is 6.3202MW by proposed method.

TABLE 5. Results of loss allocation of 12-bus radial system considered as microgrid

Bus	Electrical Closeness Centrality	Outgoing Power (in MW)	Loss allocated (in MW)
1	15.8263	111.8429	1.6618
2	10.8454	110.6288	1.1264
3	8.9293	64.7067	0.5424
4	8.9751	30.0000	0.2528
5	8.8759	90.0000	0.7500
6	8.9586	56.6397	0.4764
7	8.9201	36.3377	0.3043
8	8.9603	15.0000	0.1262
9	8.9071	60.0000	0.5017
10	8.9676	45.0362	0.3792
11	12.3343	35.0049	0.4054
12	16.8483	20.0000	0.3164
Total			6.8429

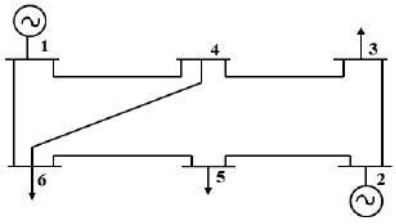


Figure 6. Six bus test system

TABLE 6. Loss allocation for 6-bus system at given load without wind generator

Bus No.	Pro-rata[2]	Z-bus [7]	Game theory	Abdelkader [8]	Elmitwally [9]	Proposed
Generators						
Bus 1	3.7252	3.4677	1.4325	0.0596	-0.1679	3.8489
Bus 2	2.5528	7.0994	4.9906	0.0498	0.4013	2.3871
Loads						
Bus 3	2.5577	0.9756	2.7365	4.8611	4.8611	2.5172
Bus 4	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Bus 5	1.3951	0.2769	2.6436	3.2991	3.2991	1.5048
Bus 6	2.3252	0.7365	0.7259	4.1624	4.1624	2.2982
Network loss	0.0000	0.0000	0.0000	0.1240	0.0000	0.0000
Total	12.5561	12.5561	12.5561	12.5561	12.5561	12.5561

4. 3. 2. 17-bus System with Radial Topology A 17-bus radial distribution network with DGs is considered as microgrid and is shown in Figure 7. The power flow results with system data is listed in Table 7 for 17-bus system. Total active power loss of the system is 6.627kW.

Proposed method is applied to 17-bus radial topology network for comparing results with some existing methods. The system data and power flow results of 17-bus radial network is similar to that of [14] and presented in Table 7. The results of comparison with few existing methods for loss contribution by loads and generators in 17-bus radial system are presented in Table 8.

From Table 8, it is clear that total loss allocated to generators is 6.627 kW while total loss allocated to loads is 3.4272 kW in 17-bus radial network. It can be seen from Table 8 that total loss allocated to loads and generators is -0.41 kW and 7.04 kW respectively by marginal method. Total loss allocated to loads is greater

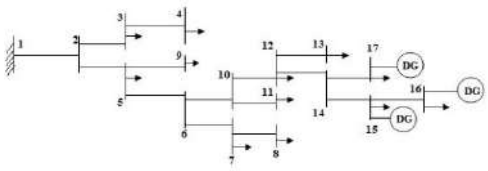


Figure 7. 17-bus radial distribution network

than that of generators by methods of [14], Z-bus and succinct. Proposed method does not allocate negative loss contributions to any participant like marginal, Z- bus or succinct methods in case of 17-bus radial network also.

From the results listed in Tables 3-8, we can see that higher amount of loss is allocated to buses having higher load or generation as expected. From the results it can be observed that, loss allocated to each bus is positive

TABLE 7. 17-bus system data [14]

Branch No.	From node	To node	r (pu)*10^{-3}	x (pu)*10^{-3}	b (pu)	From node injection (kW)	From node injection (kVAR)	To node injection (kW)	To node injection (kVAR)
1	1	2	2.5	2.6	0.03	1100.62	561.059	-1096.7	-586.919
2	2	5	0.7	0.7	0.02	896.716	493.692	-895.960	-512.750
3	2	3	0.8	0.8	0.02	200.050	93.2270	-200.010	-113.010
4	3	4	0.7	0.7	0.00	111.010	63.0120	-111.000	-63.0000
5	5	9	2.1	2.2	0.02	89.0200	30.2350	-89.0000	-50.0000
6	5	6	2.0	2.1	0.02	666.949	402.521	-665.706	-420.960
7	6	10	0.1	0.1	0.00	186.170	168.100	-186.170	-168.100
8	6	7	0.9	0.9	0.00	479.529	252.855	-479.258	-262.429
9	7	8	1.7	1.7	0.00	338.258	182.429	-338.000	-192.000
10	10	11	0.6	0.6	0.00	152.019	86.0190	-152.000	-86.0000
11	10	12	1.8	1.8	0.00	34.1520	82.0840	-34.1370	-82.0690
12	12	13	0.3	0.3	0.00	10.0000	5.0000	-10.0000	-5.0000
13	12	14	1.1	1.1	0.00	-241.860	-73.9310	241.930	74.0020
14	14	15	1.1	1.1	0.00	-222.930	-85.0840	222.990	85.1480
15	14	17	0.7	0.7	0.00	-19.0000	11.0820	19.0000	-11.0820
16	15	16	0.1	0.1	0.00	-127.998	-55.8580	128.000	55.8600

TABLE 8. Loss allocation results with load and DG data for 17-bus radial system

Bus No.	P (kW)	Q (kVAR)	Loss Allocation in kW					
			Pro rata	Marginal	Z-bus	Succinct	Jahromi [14]	Proposed
Loads								
3	89.000	50.0000	0.1600	0.3400	0.220	0.2200	0.0900	0.1995
4	111.00	63.0000	0.2000	0.4900	0.290	0.2900	0.1800	0.1186
5	140.00	80.0000	0.2500	0.5600	0.430	0.4400	0.2600	0.8285
7	141.00	80.0000	0.2500	0.5200	0.780	0.7900	0.8600	0.4778
8	338.00	192.000	0.6000	2.0800	2.120	2.1500	3.3900	0.6311
9	89.000	50.0000	0.1600	0.4900	0.300	0.3000	0.1600	0.0951
11	152.00	86.0000	0.2700	0.3700	0.770	0.7900	1.2300	0.1624
12	266.00	151.000	0.4800	-0.3000	1.360	1.4200	0.3900	0.2555
13	10.000	5.0000	0.0200	-0.0300	0.050	0.0500	0.0100	0.0107
15	205.00	116.000	0.3700	-1.9800	-0.250	-0.1500	0.0000	0.4267
16	72.000	41.0000	0.1300	-0.8000	-0.090	-0.0600	0.0000	0.2137
17	241.00	137.000	0.4300	-2.1400	-0.510	-0.1800	0.0000	0.2777
Subtotal			3.3100	-0.4100	5.470	6.0500	6.5700	3.4272
Generators								
15	300.00	145.290	1.3100	2.7900	0.360	0.2200	0.0300	1.2631
16	200.00	96.8600	0.8700	2.1000	0.240	0.1700	0.0200	0.8421
17	260.00	125.920	1.1300	2.1500	0.550	0.1900	0.0000	1.0947
Subtotal			3.3100	7.0400	1.150	0.5800	0.0500	3.1998
Total			6.6200	6.6300	6.620	6.6300	6.6200	6.6270

always which indicates that the proposed method does not allocate negative loss contribution to renewable energy resources which are DGs connected in a microgrid.

From Table 8, it can be observed that loss allocated to loads is 3.31 kW, 5.47 kW and 6.05 kW by pro-rata, Z-bus and succinct method respectively while marginal method allocates negative loss (-0.41 kW) to loads. Loss allocated to generators is 3.31 kW, 1.15 kW and 0.58kW by pro-rata, Z-bus and succinct method respectively while marginal method while marginal method makes over recovery by allocating 7.04kW loss to generators.

5. APPLICATION OF PROPOSED LOSS ALLOCATION METHOD IN LARGER MICROGRID

Consider a modified 30-bus system as interconnected microgrid with meshed topology in which distributed energy resources (DERs) are connected at five buses. Bus 13, bus 23 and bus 27 have solar plants of 30 MW, 20 MW and 30 MW rated capacity respectively. Bus 2 and bus 22 are connected with a wind farm of 60 MW and 20 MW rated capacity respectively. After power flow analysis, total system loss is 5.6436 MW. Results of LA to various buses of 30-bus microgrid system by proposed method are listed in Table 9.

TABLE 9. Results of loss allocation of 30-bus system with 5 DERs considered as microgrid

Bus	Electrical Closeness Centrality (C_i)	Outgoing Power at bus (P_i), (in MW)	Proportional Index ($a_i = \frac{P_i}{\sum_{i=1}^{n} P_i C_i}$)	Loss Allocated (in MW)
1	26.6132	93.8436	0.0047	0.7008
2	24.4788	115.9418	0.0058	0.7964
3	26.9713	36.4975	0.0018	0.2762
4	24.8902	62.6927	0.0031	0.4379
5	26.9121	26.6015	0.0013	0.2009
6	22.9696	82.7536	0.0041	0.5334
7	26.9586	22.8000	0.0011	0.1725
8	27.9586	30.0000	0.0015	0.2354
9	25.9417	20.9345	0.0010	0.1524
10	24.9225	12.5812	0.0006	0.0880
11	27.9530	10.0000	0.0005	0.0784
12	23.8812	39.1963	0.0019	0.2627
13	27.9776	30.0000	0.0015	0.2355
14	26.9403	6.2000	0.0003	0.0469
15	24.9361	18.6747	0.0009	0.1307
16	26.9189	10.8659	0.0005	0.0821
17	26.9384	9.0000	0.0004	0.0680
18	26.9236	9.9101	0.0005	0.0749
19	26.9171	9.5000	0.0005	0.0718
20	26.9218	5.0267	0.0002	0.0380
21	26.9371	18.2619	0.0009	0.1380
22	26.8745	18.3642	0.0009	0.1385
23	26.8853	20.0000	0.0010	0.1509
24	25.9338	10.0965	0.0005	0.0735
25	25.9366	14.6109	0.0007	0.1063
26	27.9202	3.5000	0.0002	0.0274
27	24.8803	15.1454	0.0008	0.1057
28	26.9485	12.4682	0.0006	0.0943
29	26.9459	6.0838	0.0003	0.0460
30	26.9459	10.6000	0.0005	0.0802
Total				5.6436

In modified 30-bus microgrid system, DGs are connected at buses 2, 13, 22, 23 and 27. From the results listed in Table 9, it can be calculated that total loss allocated to generators and loads is 1.8955 MW and 3.7482 MW respectively. It is clear that distribution of system loss among generators and loads is not like pro-rata method. The proposed method segregates loss on the basis of electrical closeness centrality which includes the relative position for sustaining transparency.

6. CONCLUSION

This paper proposes a simple and robust method for LA in interconnected microgrid with meshed topology as well as radial topology where generations are considered to be variable. Due to various non-conventional energy sources in microgrid, power flow changes very frequently and presence of multiple sources modifies network usages. Proposed method easily determines loss allocated to each bus for every scenario. While allocating loss to different buses, the proposed method considers relative position of buses in the network and requires only power flow solution with network data. Electrical closeness centrality measure is used to identify the relative location of buses present in the network and can be calculated by using power flow results. The results

obtained by proposed method shows that electrical closeness centrality changes according to power flow results and corresponding to that amount of loss allocated to different buses changes. LA by proposed method includes position of each load and generators in the network and loss contribution to generators and loads depend on the amount of power produced or consumed by them. The proposed procedure is simple to understand and its execution is undemanding because it does not require intricate computational application.

7. REFERENCES

1. Chakraborty, P., Baeyens, E. and Khargonekar, P.P., "Distributed control of flexible demand using proportional allocation mechanism in a smart grid: Game theoretic interaction and price of anarchy", ***Sustainable Energy, Grids and Networks***, Vol. 12, (2017), 30-39, doi: 10.1016/j.segan.2017.09.002.
2. Conejo, A., Arroyo, J., Alguacil, N. and Guijarro, A., "Transmission loss allocation: A comparison of different practical algorithms", ***IEEE Transactions on Power Systems***, Vol. 17, No. 3, (2002), 571-576, doi: 10.1109/TPWRS.2002.800894.
3. Zhang, S., Chen, F., Wang, X. and Zhong, J., "Study on loss allocation method of distribution network with considering of distributed generation", in 2020 IEEE/IAS Industrial and Commercial Power System Asia (I&CPS Asia), IEEE. (2020), 508-512.
4. Moret, F., Tosatto, A., Baroche, T. and Pinson, P., "Loss allocation in joint transmission and distribution peer-to-peer markets", ***IEEE Transactions on Power Systems***, Vol. 36, No. 3, (2020), 1833-1842, doi: 10.1109/TPWRS.2020.3025391.
5. Shimohammadi, D., Gribik, P., Law, E., Malinowski, J. and EO'Donnell, R., "Evaluation of transmission network capacity use for wheeling transaction", ***IEEE Transactions on Power Systs.***, Vol. 4, No. 4, (1989), 1405-1413, doi: 10.1109/59.41691.
6. Carpaneto, E., Chicco, G. and Akilimali, J.S., "Computational aspects of the marginal loss allocation methods for distribution systems with distributed generation", in MELECON 2006-2006 IEEE Mediterranean Electrotechnical Conference, IEEE. (2006), 1028-1031.
7. Mutale, J., Strbac, G., Curcic, S. and Jenkins, N., "Allocation of losses in distribution systems with embedded generation", ***IEE Proceedings-Generation, Transmission and Distribution***, Vol. 147, No. 1, (2000), 7-14, doi: 10.1109/TPWRS.2003.820698.
8. Strbac, G., Kirschen, D. and Ahmed, S., "Allocating transmission system usage on the basis of traceable contributions of generators and loads to flows", ***IEEE Transactions on Power Systems***, Vol. 13, No. 2, (1998), 527-534, doi: 10.1109/59.667378.
9. Conejo, A.J., Galiana, F.D. and Kockar, I., "Z-bus loss allocation", ***IEEE Transactions on Power Systems***, Vol. 16, No. 1, (2001), 105-110, doi: 10.1109/59.910787.
10. Abdelkader, S.M., "Characterization of transmission losses", ***IEEE Transactions on Power Systems***, Vol. 26, No. 1, (2010), 392-400, doi: 10.1109/TPWRS.2010.2052115.
11. Elmitwally, A., Eladl, A. and Abdelkader, S.M., "Efficient algorithm for transmission system energy loss allocation considering multilateral contracts and load variation", ***IET Generation, Transmission & Distribution***, Vol. 9, No. 16, (2015), 2653-2663, doi: 10.1049/iet-gtd.2015.0597.
12. Carpaneto, E., Chicco, G. and Akilimali, J.S., "Characterization of the loss allocation techniques for radial systems with distributed generation", ***Electric Power Systems Research***, Vol. 78, No. 8, (2008), 1396-1406, doi: 10.1016/j.epsr.2008.01.001.
13. Atanasovski, M. and Taleski, R., "Power summation method for loss allocation in radial distribution networks with dg", ***IEEE transactions on Power Systems***, Vol. 26, No. 4, (2011), 2491-2499, doi: 10.1109/TPWRS.2011.2153216.
14. Atanasovski, M. and Taleski, R., "Energy summation method for loss allocation in radial distribution networks with dg", ***IEEE transactions on Power Systems***, Vol. 27, No. 3, (2012), 1433-1440, doi: 10.1109/TPWRS.2011.2182663.
15. Savier, J. and Das, D., "An exact method for loss allocation in radial distribution systems", ***International Journal of Electrical Power & Energy Systems***, Vol. 36, No. 1, (2012), 100-106, doi: 10.1016/j.ijepes.2011.10.030.
16. Ghofrani-Jahromi, Z., Mahmoodzadeh, Z. and Ehsan, M., "Distribution loss allocation for radial systems including dgs", ***IEEE Transactions on Power Delivery***, Vol. 29, No. 1, (2013), 72-80, doi: 10.1109/TPWRD.2013.2277717.
17. Jagtap, K.M. and Khatod, D.K., "Loss allocation in radial distribution networks with various distributed generation and load models", ***International Journal of Electrical Power & Energy Systems***, Vol. 75, (2016), 173-186, doi: 10.1016/j.ijepes.2015.07.042.
18. Kumar, P., Gupta, N., Niazi, K.R. and Swarnkar, A., "A circuit theory-based loss allocation method for active distribution systems", ***IEEE Transactions on Smart Grid***, Vol. 10, No. 1, (2017), 1005-1012, doi: 10.1109/TSG.2017.2757059.
19. Sharma, S. and Abhyankar, A., "Loss allocation for weakly meshed distribution system using analytical formulation of shapley value", ***IEEE Transactions on Power Systems***, Vol. 32, No. 2, (2016), 1369-1377, doi: 10.1109/TPWRS.2016.2571980.
20. Bharti, D. and De, M., "A centrality index based approach for selection of optimal location of static reactive power compensator", ***Electric Power Components and Systems***, Vol. 46, No. 8, (2018), 886-899, doi: 10.1080/15325008.2018.1514548.

Persian Abstract

چکیده

با برخی فرض ها و محدودیت ها ، روش های مختلفی در ادبیات به طور عمده برای تخصیص تلفات در شبکه انتقال و پس از آن برای شبکه توزیع شعاعی گسترش یافته و برخی از روش ها به طور خاص برای شبکه توزیع شعاعی توسعه یافته است. اما این روش ها برای میکرو شبکه هایی که با شبکه های معمولی در سطح انتقال و توزیع فرعی بسته به موقعیت جغرافیایی یکپارچه هستند ، مناسب نیستند. در این مقاله یک روش تخصیص تلفات بر اساس نتایج جریان برق و موقعیت نسبی اتوبوس ها برای میکرو شبکه بهم پیوسته ارائه می شود که در صورت تغییر مکرر نسل ها به دلیل تناوب منابع تجدید پذیر بسیار موثر است. اجرای روش پیشنهادی در ریز شبکه با توپولوژی مشبک و شعاعی بدون هیچ گونه پیچیدگی محاسباتی بسیار ساده است و فقط به راه حل جریان برق و داده های شبکه نیاز دارد. نتایج برای شرایط مختلف تولید منابع تجدید پذیر در ریز شبکه نشان داده شده است تا کارآیی روش تخصیص تلفات پیشنهادی را نشان دهد.

International Journal of Engineering

Journal Homepage: www.ije.ir

A New Simple Method to Avoid Maloperation of Transformer Restricted Earth Fault Relay during External Fault Events

A. Ebadi, S. M. Hosseini*, A. A. Abdoos

Department of Electrical and Computer Engineering, Babol Noshirvani, University of Technology, Babol, Iran

P A P E R I N F O

Paper history:
Received 28 April 2021
Received in revised form 09 June 2021
Accepted 28 July 2021

Keywords:
Power Transformer Protection
Restricted Earth Fault
Jiles-Atherton Model
Current Transformer Saturation

A B S T R A C T

The restricted earth fault (REF) protection is provided for electrical power transformer in order to sense internal earth faults, mainly because it is more sensitive than the main differential protection. However, The REF relay may maloperate when current transformer (CT) saturation happens following a severe external fault. In this paper, a new simple algorithm is proposed for REF protection scheme which is realized by considering four fundamental conditions. These conditions are defined based on the differential and neutral currents as well as sum of phase currents. When these conditions are simultaneously satisfied, the relay detects the internal ground fault. This algorithm is implemented and evaluated by MATLAB program based on obtained result data from simulation of a real power system using PSCAD/EMTDC software package. The well-known Jiles-Atherton (JA) model is used to simulate the transient behavior of CTs. The satisfactory results obtained from exhaustive investigation justify the high security of the proposed protection scheme.

***doi**: 10.5829/ije.2021.34.09c.04*

NOMENCLATURE

I_d	Differential current	I_{sum}	Residual current
I_n	Neutral current	I_d^{TH}	Threshold value of differential current
I_a, I_b, I_c	Phase currents	I_n^{TH}	Threshold value of neutral current

1. INTRODUCTION

Power transformers are the most important and expensive equipment of high voltage substations. Thus, they should be protected against any internal faults by sensitive protective relays. Differential protection is the most commonly used transformer protection scheme. It can cover most of short circuits inside the transformer protection zone including terminal and winding faults [1]. However, when an earth fault happens close to neutral point of the transformer Y winding or the fault impedance is high, the fault current may not be sufficient to drive the differential protection scheme. The restricted earth fault (REF) protection scheme is designed to respond to such earth faults [2]. But, current transformer (CT) saturation caused by either the magnetizing inrush currents or severe external faults may make the REF relay maloperate [3-5]. It should be noted that the strictly CT dimensioning can improve the protection system immunity against CT saturation; however, the issue is not solved completely [6]. Hence, some methods have been proposed to make this protection scheme stable during inrush current and external faults. The first alternative was a high-impedance REF prtection scheme, but this method needs phase and neutral CTs with exactly the same magnetizing characteristic as well as a high knee-point voltage [7, 8]. Nowadays, numerical low-impedance REF relays have received more attention because they do not need the rigid CT requirements of the high-impedance type [9]. However, during CT saturation, a false differential current appears in the low-impedance REF relay and this may cause the relay maloperation [10-12]. Adaptive restraint currents and directional supervision methods can improve this relay

*Corresponding Author Institutional Email: mehdi.hosseini@nit.ac.ir (S. M. Hosseini)

Please cite this article as: A. Ebadi, S. M. Hosseini, A. A. Abdoos, A New Simple Method to Avoid Maloperation of Transformer Restricted Earth Fault Relay during External Fault Events, International Journal of Engineering, Transactions C: Aspects, Vol. 34, No. 09, (2021) 2070-2076

performance but even with these methods, the maloperation issue remains in many cases [13]. A research shows that by combining the REF and earth fault (EF) units, the protection stability against maloperation improves to some extent [14]. However, if the neutral current is enough high during a potential case of maloperation, this method will not be helpful. The Krstivojevic and Djurić [15, 16] proposed a new REF algorithm based on a time domain phase comparator that yields desirable results for inrush current condition and only single phase to ground type of external fault. Ebadi et al. [17] have designed a new REF relay based on artificial intelligence which is slower than conventional method. They also employed time-frequency analysis to create an effective supervision method for conventional REF relays [18].

Based on above descriptions, more efforts must be done to tackle this issue. In this paper, a new algorithm is proposed that significantly improves the REF protection immunity during external faults. Furthermore, during an internal fault event, the neutral current vector is in opposite direction of the residual current vector. On the other hand, they have the same direction during an external fault event. So, the magnitude of differential vector of these two currents is calculated using different formulas for internal and external faults. In this study, four different conditions are defined based on the calculated differential current vector. The proposed REF protection scheme issues the trip signal when all defined conditions are simultaneously satisfied. The logic behind the algorithm is very simple and it can be easily implemented in practical applications. Moreover, it is very fast to detect internal faults.

The rest of this paper is organized as follows: In section 2, the test study is introduced. Section 3 presents the proposed method. Simulation results are presented in section 4. Finally, the conclusion is made at the end of the paper in section 5.

2. TEST STUDY

For performance evaluation of the proposed method in internal and external fault events, a part of the Iranian high voltage substation depicted in Figure 1 is simulated using PSCAD/EMTDC software package. This 50 Hz power system includes a 230/63 kV power transformer with nominal power of 160 MVA, a grounding transformer and CTs with turns ratios of 600:1 and 2000:1 installed at the Y and Δ sides of the power transformer, respectively. More details about the power system under study have been presented by Ebadi et al. [18]. It is notable that the precise Jiles–Atherton model is applied to simulate CTs. The parameters of Jiles–Atherton model according to magnetization characteristics of CTs can be found in literature [19].

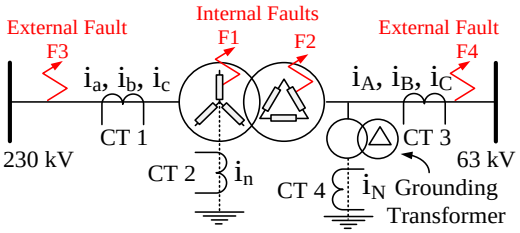


Figure 1. The single line diagram of power system under study

3. PROPOSED METHOD

A typical differential (operating) current of low-impedance REF relays is given by Equation (1):

$$I_d = (I_a + I_b + I_c) - I_n = I_{sum} - I_n \quad (1)$$

where I_a, I_b and I_c denote phase currents, and I_n is the neutral current. I_{sum} Indicates the sum of phase currents that is called "residual current".

Figure 2(a) shows that the neutral and residual currents flow in the same direction (according to the assumed direction of the currents) during external fault. Thus, the magnitude of differential current vector can be calculated using Equation (2):

$$|I_d| = \big||I_{sum}| - |I_n|\big| \quad (2)$$

On the other hand, the neutral and residual currents flow in opposite direction during internal fault as can be seen in Figure 2(b). Thus, according to expression Equation (3), the magnitude of differential current vector is equal to the sum of magnitudes of the neutral and residual current vectors, in this case. It must be noted that if the circuit of faulty side of power transformer is opened and it is energized from other side, the magnitude of differential current will be equal to neutral current (see Figure 3(c)).

$$|\mathrm{I_d}| = |\mathrm{I_{sum}}| + |\mathrm{I_n}| \quad (3)$$

From the comparison of expressions Equations (2) and (3), it can be concluded that the following constrains are met only in internal fault condition. Thus, these constraints can be used to detect internal fault from external fault.

$$|\mathrm{I_d}| > |\mathrm{I_{sum}}| \text{ and } |\mathrm{I_d}| \geq |\mathrm{I_n}| \quad (4)$$

Therefore, to diagnose an abnormal condition that may origin from an internal ground fault, the differential and neutral currents can be compared to predetermined threshold values as follow expression. Then, the internal ground fault can be detected using Equation (4). Consequently, the new detection mechanism of REF relay is designed as shown in Figure 3.

$$|\mathrm{I_d}| \geq \mathrm{I_d^{TH}} \text{ and } |\mathrm{I_n}| \geq \mathrm{I_n^{TH}} \quad (5)$$

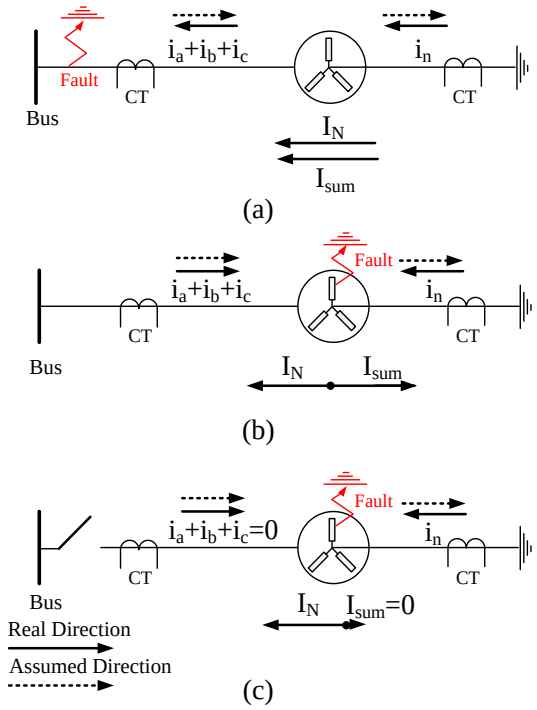


Figure 2. The Y side of power transformer during external fault (a), internal fault with non-zero phase currents (b) and internal fault with zero phase currents (c)

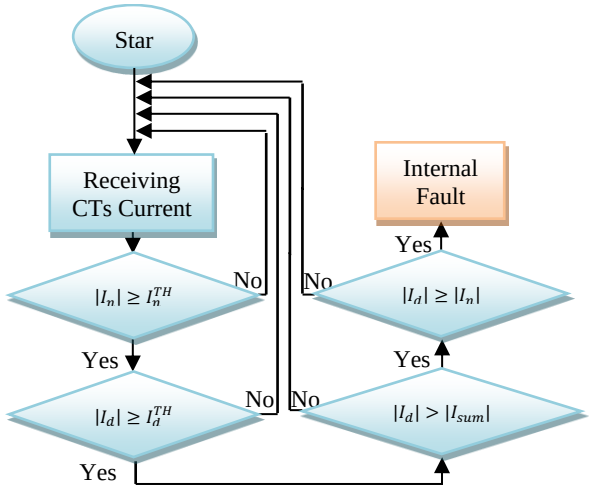


Figure 3. The proposed algorithm for detecting internal ground fault from external faults

4. SIMULATION RESULTS

In this section, initially, the performance of the proposed method is compared to a conventional method with recommended setting [20] for a number of fault events. It must be noted, I_d^{TH} is set to 10% which is the minimum value for differential current in REF relay setting [20]. Also, 1% as a very small value is chosen for I_n^{TH}. By applying these settings, a strict evaluation of the proposed method performance is achieved. Then, the detection accuracy of the proposed method is evaluated for a wide range of internal and external fault events in comparison with the conventional method and an intelligent method [17]. It must be noted that two methods are simulated using MATLAB program but there is no interface between this program and PSCAD. Furthermore, MATLAB processes only saved simulation data obtianed using PSCAD.

4. 1. Performance Evaluation during Some Sample Fault Events In this subsection, three different fault scenarios are considered to study the performance of the proposed and conventional methods. The first scenario is a severe double line to ground external fault with zero fault resistance which occurs on the Δ side of the power transformer at time=100 ms, while it is connected to the voltage source from the Y side and supplies a full-load power. The second scenario is a severe single line to ground internal fault with zero fault resistance which occurs on the Y side terminal of the power transformer at time=100 ms. The third scenario is a light internal fault similar to second scenario with 3 kΩ of fault resistance. Also, the voltage source and load of two last scenarios are similar as the first scenario.

Figure 4 shows that despite that the differential current is zero during the external fault assuming use of unsaturable (ideal) CTs; however, a high magnitude spurious differential current appears using real CTs. Thus, according to Figure 5, the current trajectory calculated based on the conventional method inserts into operation region (see Figure 5) and it maloperates 47.8 ms after fault inception, but the new method remains stable (see Figure 6)

The differential current for the severe internal fault is depicted in Figure 7. It is seen that the differential current waveform obtained from real CTs is deeply distorted as compared to real CTs. According to Figure 8, the current trajectory of the conventional method inserts quickly into

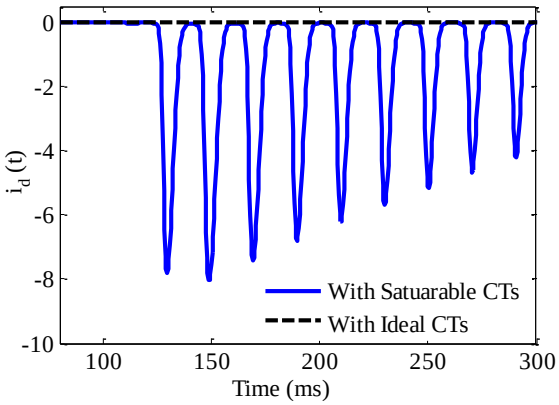


Figure 4. Differential current during the external fault

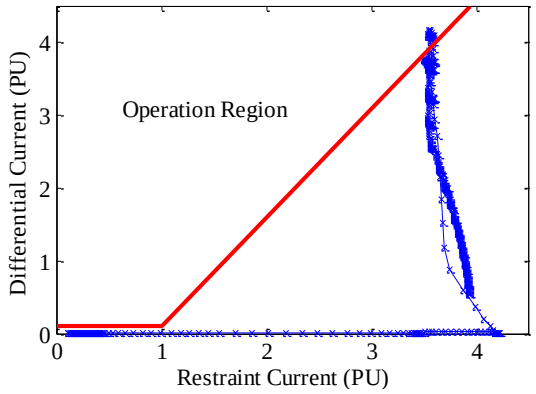


Figure 5. Caculated current trajectory for the external fault based on the conventional method

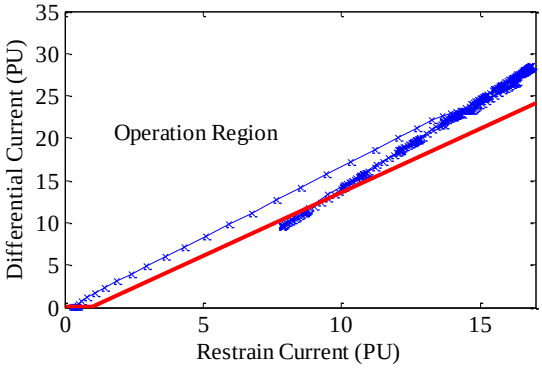


Figure 8. Calculated current trajectory for the severe internal fault based on the conventional method

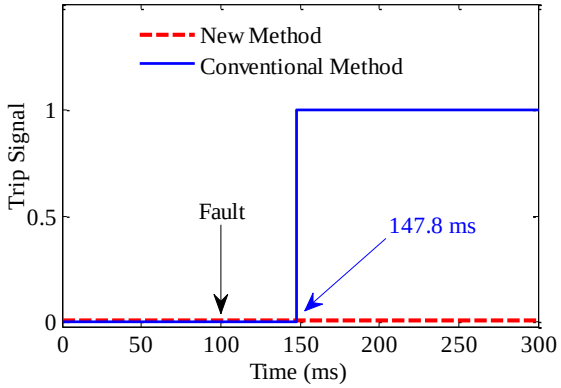


Figure 6. Trip signals of the proposed and conventional methods for the external fault

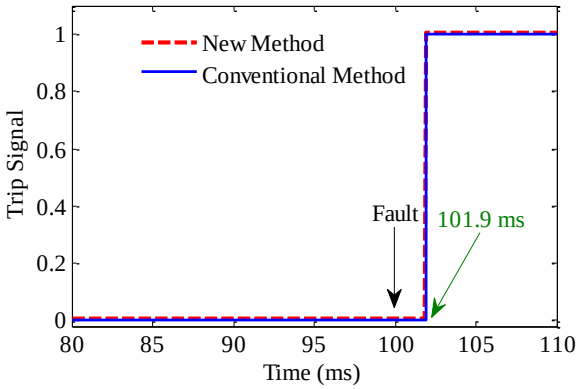


Figure 9. Trip signals of the proposed and conventional methods for the severe internal fault

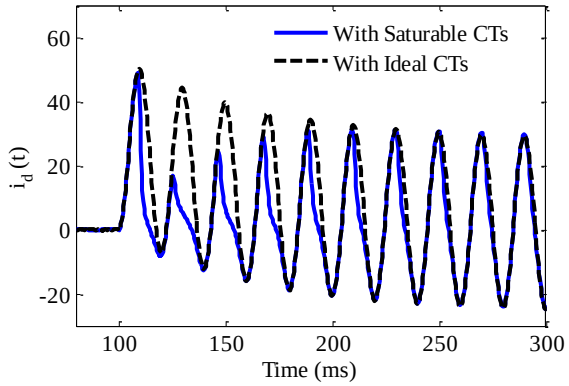


Figure 7. Differential current during the severe internal fault

the operation without any issue. The conventional method needs only 1.9 ms to detect the fault, exactly similar to new method (see Figure 9). It means that the fault detection time is less than one cycle, for this severe fault case.

Figure 10 indicates the differential current during the light internal fault. As can be seen in Figure 11, the current trajectory computed based on the modeled conventional method inserts into the operation region. The time taken by the conventional and the new method for detection of the fault is 17 ms (see Figure 12) that is still less than one cycle.

4. 2. A Comprehensive Comparative Accuracy Assessment To obtain a comprehensive and comparative assessment of the proposed method

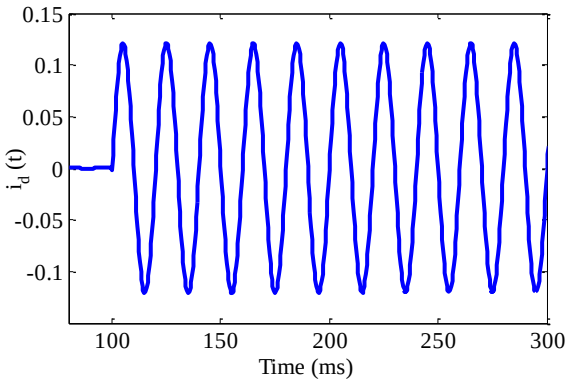


Figure 10. Differential current during the light internal fault

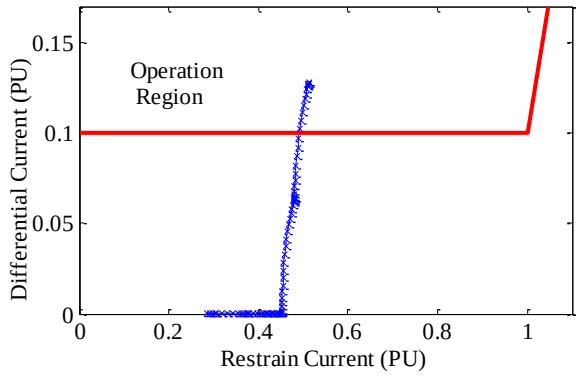


Figure 11. Calculated current trajectory for the light internal fault based on the conventional method

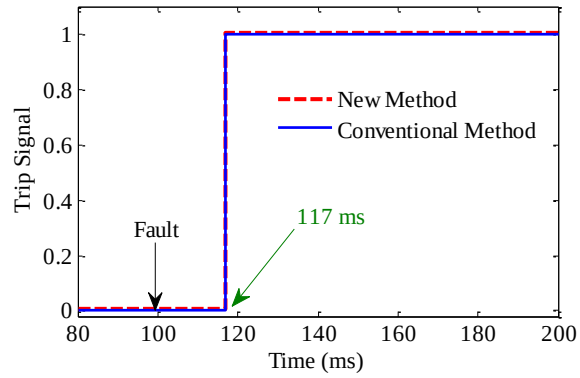


Figure 12. Trip signals of the proposed and conventional methods for the light internal fault

detection accuracy, its performance is evaluated for a wide range of internal and external faults in comparison with the conventional method [20] and the intelligent method [17].

To simulate internal fault events, voltage source is connected to the Y winding of power transformer under study and it supplies full-load power. For fault locations, 10 points on the one phase of Y winding and 11 points on the one phase of Δ winding are chosen. These 10 points is placed at distances equal to 0%, 10%, …, 90% of winding length from the terminal end. Also, the mentioned 11 points are distributed uniformly on the winding so that the first is at the middle of the winding and the last is at the end of winding connecting to the output terminal. Besides, three different values are considered for the fault resistance. The lowest value is 0 Ω, the highest is set so that to obtain the lightest fault which is detectable by the conventional method (I_d=10 %) and the third value is set to a value between them. Also, to take into account the fault inception angle effects, 11 points distributed uniformly over a cycle are chosen as fault occurrence time. On the other hand, the remnant flux of CTs are selected from -85%, 0% and +85% of rated flux. All in all, 1053 internal ground fault cases are simulated for this part of study.

Besides, 880 cases of external faults are simulated by considering a large number of single line-to-ground faults, line-to-line faults, double line-to-ground faults and three-phase faults. In order to precisely evaluate the proposed method performance, very severe faults are simulated by considering low values for fault resistances. On the other hand, external faults can be categorized into two scenarios. In the first scenario, the faults occur on the Y side (F3) while the voltage source is connected to the Δ winding of the power transformer at full load condition. In the next scenario, the faults occur on the Δ side (F4) while the power transformer is fully loaded and the Y winding is connected to the voltage supply. It is worth noting that various values for fault inception time and residual flux of CTs are considered in the simulations, like as internal fault cases.

Based on obtained results from simulated test case in the PSCAD/EMTDC software, three mentioned methods have been implemented and evaluated in the MATLAB environment. As given in Tabel 1, it can be seen that all internal faults have been detected by three methods without any issue. However, the conventional and intelligent methods could not remain stable during 203 and 13 external fault cases, respectively. Besides, the proposed method has maloperation for only 9 external faults. It means that the proposed method with total accuracy of 99.53% has the best performance compared to the conventional and intelligent methods with accuracy of 88.1% and 99.32%, respectively. It can be concluded that that the high accuracy of the proposed method as well as its simplicity make it possible for practical implementation.

TABLE 1. Comparison results of different methods

Method	# of misoperation for 1053 cases of internal fault	# of maloperation for 880 cases of external fault	Total Accuracy (%)
Conventional Method [20]	0	203	88.10
Intelligent Method [17]	0	13	99.32
Proposed Method	0	9	**99.53**

5. CONCLUSION

A restricted earth fault relay may maloperate during severe external fault events due to CT saturation. To tackle this issue, this paper presents a new simple algorithm which has high stability in such situation. To evaluate the proposed protection scheme performance, a real power system has been simulated using PSCAD/EMTDC program considering a huge number of

internal and external faults conditions. In order to show the superiority of the proposed method, its performance has been compared with a conventional and an intelligent REF relay. This study proved that the proposed method, in spite of its simplicity, can accurately detect internal faults from external faults while it is as fast as the conventional method. Since the REF relay should be immune against maloperation during inrush current in addition to the external fault, a new method for this purpose can be developed as a future work.

6. AKNOWLEDGEMENT

The authors gratefully acknowledge the funding support of Babol Noshirvani University of Technology with grant program numbers BNUT/370199/00 and BNUT/390066/00.

7. REFERENCES

1. Ganji, M., Bigdeli, M. and D. Azizian, D., "Mitigation Transformer Inrush Current using Modified Transient Current Limiter", ***International Journal of Engineering, Transactions B: Applications***, Vol. 32, No. 5, (2019), 701-709. DOI: 10.5829/ije.2019.32.05b.12
2. Ebadi, A., Hosseini, S.M. and Abdoos, A.A., "Immunity Improvement of a Low-Impedance Restricted Earth Fault Relay by Making its Mechanism Intelligent Based on Support Vector Machine", ***Computational Intelligence in Electrical Engineering***, Vol. 9, No. 2, (2018), 29-40. DOI: 10.22108/isee.2018.111464.1132
3. Ebadi, A., Hosseini, S.M. and Abdoos, A.A., "Designing of a New Transformer Ground Differential Relay Based on Probabilistic Neural Network", ***Energy Engineering & Management***, Vol. 9, No. 4, (2020), 2-13. DOI: 10.22052/9.4.2
4. Krstivojevic, J.P. and Djurić, M.B., "New Algorithm for Transformer Ground Fault Protection", in MedPower 2014 (IET), Athens, (2014), 1-6. DOI: 10.1049/cp.2014.1698
5. Dinesh Babu, K.N., Ramaprabha, V. Rajini, V. and Nagarajan, V., "A Case Study on REF Low Impedance IED Mal Operation", in IEEE 6th International Conference on Power Systems, New Delhi, (2016), 1-5. DOI: 10.1109/ICPES.2016.7583997
6. Taghipour Gorji, R., Hosseini, S. M., Abdoos, A. A. and Ebadi, A., "A Hybrid Intelligent Method for Compensation of Current Transformers Saturation Based on PSO-SVR", ***Periodica Polytechnica Electrical Engineering and Computer Science***, Vol. 65, No. 1, (2021), 53-61. DOI: 10.3311/PPee.16248
7. Apostolopoulos, C. And Tsakiris, D., "Design And Performance Evaluation of a High-Impedance REF Scheme for MV/LV Transformers", ***IEEE Transactions on Industry Applications***, Vol. 51, No. 6, (2015), 5398-5409. DOI: 10.1109/TIA.2015.2435003
8. Subramanian, P.V. and Ajithal, L., "Practical Considerations for High Impedance Restricted Earth Fault Relay Settings", ***Journal of the Institution of Engineers (India): Series B***, Vol. 96, No. 1, (2015), 107-110. DOI: 10.1007/s40031-014-0124-x
9. Ebadi, A., Hosseini, S.M. and Abdoos, A.A., "A New Artificial Intelligence based Supervision Method for Low-Impedance REF Relays", ***Electric Power Systems Research***, Vol. 195, (2021), 107177. DOI: 10.1016/j.epsr.2021.107177
10. Nasution, M.F., Mustafa, F. and Shaulgara, S., "Case Studies of Magnetizing Inrush Current Effect on Differential & REF Transformer Protection", in 2nd International Conference on High Voltage Engineering and Power Systems (ICHVEPS), Bali, (2019), 1-6. DOI: 10.1109/ICHVEPS47643.2019.9011114
11. Nimtaj, B., Mahmoudi, A., Palizban, O. and Kahourzade, S., "A Comparison of Two Numerical Relay Low Impedance Restricted Earth Fault Algorithms in Power Transformer", in 8th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology, Khon Kaen, (2011), 7921-795. DOI: 10.1109/ECTICON.2011.5947959
12. Nim Taj, B., Mahmoudi, A. and Kahourzade, S., "Comparison of Low-Impedance Restricted Earth Fault Protection in Power Transformer Numerical Relays", ***Australian Journal of Basic and Applied Sciences***, Vol. 5, No. 12, (2011), 2458-2474.
13. Aires, M.N.O., Medeiros, R.P., Costa, F.B., Silva, K.M., Chavez, J.J. and Popov, M., "A wavelet-based restricted earth-fault power transformer differential protection", ***Electric Power Systems Research***, Vol. 196, (2021), 107246. DOI: 10.1016/j.epsr.2021.107246.
14. Davarpanah, M., M. Sanaye-Pasand, M. Iravani, R., "Performance Enhancement of the Transformer Restricted Earth Fault Relay", ***IEEE Transactions on Power Delivery***, Vol. 18, No. 1, (2012), 467-474. DOI: 10.1109/TPWRD.2012.2208204
15. Krstivojevic, J.P. and Djurić, M.B., "A New Method of Improving Transformer Restricted Earth Fault Protection", ***Advances in Electrical and Computer Engineering***, Vol. 14, No. 3, (2014), 41-48. DOI: 10.4316/AECE.2014.03005
16. Krstivojevic, J.P. and Djurić, M.B., "A New Algorithm for Avoiding Maloperation of Transformer Restricted Earth Fault Protection Caused by The Transformer Magnetizing Inrush Current And Current Transformer Saturation", ***Turkish Journal of Electrical Engineering and Computer Sciences***, Vol. 24, No. 6, (2016), 5025-5042. DOI: 10.3906/elk-1409-92
17. Ebadi, A., Hosseini, S.M. and Abdoos, A.A., "A New Restricted Earth Fault Relay Based on Artificial Intelligence", ***International Journal of Engineering, Transactions A: Basics***, Vol. 32, No. 1, (2019), 69-70. DOI: 10.5829/ije.2019.32.01a.08
18. Ebadi, A., Hosseini, S.M. and Abdoos, A.A., "A New Time-Frequency Analysis Based Supervision Method for the Transformer Restricted Ground Fault Relay", ***International Journal of Electrical Power and Energy Systems***, Vol. 129, (2021), 106858. DOI: 10.1016/j.ijepes.2021.106858
19. Gil, M. and Abdoos, A.A., "Intelligent busbar protection scheme based on combination of support vector machine and S-transform", ***IET Generation, Transmission & Distribution***, Vol. 11, No. 6, (2017), 2056-2064. DOI: 10.1049/iet-gtd.2016.1686
20. MiCOM P642, P643 & P645, "Transformer protection relays", Technical Manual, (2011), 321-326.

Persian Abstract

چکیده

حفاظت خطای زمین محدود شده برای آشکارسازی خطاهای داخلی زمین ترانسفورماتور قدرت بکارگرفته می‌شود بطوریکه جهت این کار از حفاظت اصلی دیفرانسیل حساس‌تر می‌باشد. با این حال، رله خطای زمین محدود شده ممکن است بدلیل اشباع ترانسفورماتورهای جریان حین خطای خارجی عملکرد کاذب داشته باشد. در این مقاله، یک الگوریتم ساده برای این طرح حفاظتی پیشنهاد می‌گردد که بر پایه چهار شرط اساسی استوار می‌باشد. این شرایط بر مبنای جریان‌های دیفرانسیل، نول و مجموع جریان فازها تعریف می‌گردند. هنگامیکه این شرایط بطور همزمان برقرارند، بمعنای خطای داخلی زمین می‌باشد. این الگوریتم در محیط نرم‌افزار MATLAB پیاده‌سازی و ارزیابی شده و برای این کار از نتایج شبیه‌سازی یک سیستم قدرت واقعی در نرم‌افزار PSCAD/EMTDC استفاده گردیده است. مدل مشهور جیلز اثرتون برای شبیه‌سازی رفتار گذرای ترانسفوورماتورهای جریان استفاده می‌شود. نتایج حاصله امنیت بالای طرح پیشنهادی را تایید می‌نماید.

IJE TRANSACTIONS C: Aspects, Vol. 34, No. 09, (September 2021) 2077-2086

International Journal of Engineering

Journal Homepage: www.ije.ir

Modeling Traffic Signal Control System at an Isolated Intersection using Queuing Systems

M. Taheri[a], J. Arkat*[a], H. Farughi[a], M. Pirayesh[b]

[a] *Department of Industrial Engineering, University of Kurdistan, Sanandaj, Iran*
[b] *Department of Industrial Engineering, Faculty Engineering, Ferdowsi University of Mashhad, Mashhad, Iran*

P A P E R I N F O

Paper history:
Received 11 April 2021
Received in revised form 28 June 2021
Accepted 24 July 2021

Keywords:
Traffic Signal Timing
Queueing System
Isolated Intersection
Fixed-time Control System

A B S T R A C T

As the population grows in cities worldwide, the number of vehicles present on the roadways also increases, resulting in slow-moving and congested traffic. Therefore, a widespread problem in large cities concerns the traffic in the streets. Traffic signals are one of the most powerful tools available to city authorities for urban traffic control. Their proper installation can improve both traffic flow and the safety of all road users. Extensive research has been conducted to reduce the impacts of long car queues, based mainly on traffic signal timing optimization. This paper estimates the average waiting time at an isolated intersection and optimizes the timing of the green and red phases using an analysis of queueing systems. The control system is assumed to be the fixed-time type, and the Poisson process is considered for the arrivals. The proposed model is applied to real traffic data at a two-phase intersection in Bojnurd, Iran. It needs to be noted that the current situation at the intersection under study reduces average waiting time only for one side, but the analytic model can reduce average waiting time for the whole intersection. Moreover, simulation experiments are carried out, the results of which verify the capabilities of the proposed methodology in traffic signal control applications.

***doi**: 10.5829/ije.2021.34.09c.05*

NOMENCLATURE

Sets and Indices		N_i	Intersection capacity
i	Movement phase index ($i = 1,2$)	w_i	Average waiting time in phase i
Parameters		W	Average waiting time at the intersection
λ_i	Arrival rate	**Decision Variables**	
μ_i	Service rate	A_1	Green light time length for phase 1
C	Yellow light time length	B_1	Red light time length for phase 1
n_{i1}	Number of vehicles in phase i	A_2	Red light time length for phase 2
n_{i2}	Signal state for phase i	B_2	Green light time length for phase 2
n_{i3}	The current state of the green or red light in phase i.		

1. INTRODUCTION

Traffic congestion is a serious problem in urban areas, where transportation demand exceeds road capacity. A frequent aspect is the induced air pollution with negative effects on health and living environments and the global economy due to the wasted time. Mere construction of new roads might not provide the best solution to congestion problems due to the enormous financial requirements and complex network effects. Instead, there is huge potential to improve the conditions through efficient traffic management and optimization of transportation networks. There are two major aspects to analyze and optimize urban transportation networks: traffic assignment, which is an important tool for forecasting traffic flow over the urban transportation network, and traffic signal timing, used to improve the services. Traffic signals are often controlled as fixed-time

* Corresponding Author's Iinstitutional Email: j.arkat@uok.ac.ir (J. Arkat)

Please cite this article as: M. Taheri, J. Arkat, H. Farughi, M. Pirayesh, Modeling Traffic Signal Control System at an Isolated Intersection using Queuing Systems, International Journal of Engineering, Transactions C: Aspects, Vol. 34, No. 09, (2021) 2077-2086

or real-time. Each of these strategies can be subdivided into isolated-intersection (controlling a single intersection and disregarding others) and coordinated-intersection (considering more than one intersection). The latter can be further subdivided into arterial and grid-network. When several closer intersections on an arterial are independent of signal control, the upstream vehicles are likely to meet the red light at the intersection downstream. The isolated control method applied among intersections unavoidably causes frequent stops. The main characteristic of arterial control is that the same cycle is established, and there are several semaphores with relative phase differences. Arterial control is appropriate for intersections with relatively short distances and heavy traffic flow. Grid-network control is an extension of arterial control, which adopts coordinated control to several sets of semaphores on a vast area over the road network.

In the fixed-time mode, signal timing is scheduled in advance for a specified period. The fixed-time signal control uses preset time intervals repeated every time the signal cycles, regardless of traffic volume changes. Some fixed-time systems use different preset time intervals for the morning or evening rush hours and other busy times. The fixed-time control system is the simplest type, and a great deal of research has been conducted to assess the performance of intersections with this control system. One of the first studies on signal timing with the fixed-time control system is the work by Webster [1], where an analytic model was presented to set the duration of the green signal and the fixed cycle length. The objective is to minimize vehicle delay at an isolated intersection, and the arrivals are assumed to have the Poisson distribution. After that, many papers were focused on fixed-time signal control. For example, Miller [2] and Newell [3] proposed approximation approaches to calculating the residual queue length at the end of the green phase in a fixed-time control system. For the first time, Heidemann [4] presented a relationship between the distribution function of queue length and delay with Poisson-distributed arrivals, an isolated intersection, and a fixed-time control system. Hu et al. [5] proposed an $M/D^X/1$ queueing model with server vacations for a fixed-time control system, which could be considered somewhat as a generalization of Heidemann's work for cases with several lanes in each street at the intersection. Chanloha et al. [6] compared the performance of the Q-learning framework to that of the M/M/1 and D/D/1 models to signal timing at an isolated intersection with a fixed cycle length. The total delay for each queuing model and, accordingly, the optimal green time were obtained on that basis, and the results were then compared to those of the Q-learning algorithm. The results indicate that the Q-learning algorithm can significantly improve network throughput and total delay with respect to those in queueing models. Van Den et al. [7] compared a queueing model for fixed-time signal timing with a batch-service queueing model and presented new equations for average vehicle delay. Habibi et al. [8] presented two algorithms to reduce traffic density and delay. Akçelik and Rouphail [9] studied a queueing model with batch arrivals to optimize the number of vacations and average queue length at an isolated intersection with a fixed-time control system. Pacheco et al. [10] analyzed an M/D/1 queue with vacations to estimate queue length variance and delay at an intersection with a fixed-time control system. Yang and Shi [11] proposed a queueing model with batch services and batch arrivals, where the objective is to minimize the average queue length on a multilane road. Ghasemi and Rasekhi [12] proposed an approach for predicting traffic signals using game theory and neural networks with swarm particle optimization. Boon et al. [13] derived the queue length distribution for a fixed-time control system, avoiding the computational challenges that previous studies had faced to solve several characteristic equations. Sumi and Ranga [14] proposed a new intelligent traffic management system (TMS), an approach for smart cities to control traffic lights and ease ambulance movement in cities. Amini and Shahi [15] and Faghri [16] investigated the influence of geometric and control features on the quality of traffic services.

Another type of traffic signal control system is the real-time control system, in which the traffic signal timing is carried out simultaneously with the inspection of traffic status using automatic cameras, and the green and red times in each cycle depend on the intersection traffic status. The real-time traffic signal control system can be of two sub-types: actuated and adaptive traffic signal control. Unlike fixed-time control, actuated control constantly adjusts the timing of the green light and, in some cases, the order of the phases. These settings are based on the traffic demand criteria recorded by the detectors at the intersections. This control method usually reduces delay and increases capacity and can be safer than fixed-time control, but it is very expensive to implement and requires advanced training for proper execution. In an adaptive control system, the traffic signal time constantly varies by the changes in vehicle arrival patterns at the intersection. This traffic information is collected by the detectors at the intersection and then evaluated. Finally, correction is made to the signal timing, where the traffic signal times are updated.

However, not many papers have considered real-time traffic signal timing control systems. Zhang and Wang [17] investigated an actuated traffic signal control system to minimize mean vacation and mean queue length and maximize vehicle throughput. Jiao et al. [18] presented a multi-objective model to minimize average delay, minimize the average number of stops, maximize traffic capacity for an inductive traffic control system, and optimize cycle time and time-varying green time using the Particle Swarm Optimization (PSO) algorithm.

Mirchandani and Zou [19] proposed an M/G/1 queueing model for an adaptive traffic signal control system. The objective is to minimize the mean queue length and average waiting time for vehicles. In that research, an intersection with two straight movement phases is considered and turns taken to the right, and left is ignored. An adaptive control system is also considered. When the signal shows the green light at one side of the intersection, the system will not turn it into red until the queue on that street is entirely empty. Recently, Chedjou and Kyamakya [20] have reviewed the studies on traffic signal control systems, investigating various strategies and their strengths and drawbacks, along with the challenges involved in their applications.

Table 1 summarizes some articles in this field that are categorized based on the intersection type, decision variables, and control system type. Despite the extensive research on traffic signal timing, some problems still require further inspection. Since the residual queue at the end of the green time is hard to obtain precisely, most studies have used approximate equations to calculate its values in fixed-time control systems and failed to provide strict equations for the objective functions. This paper proposes an analytic approach to obtain average waiting time at an isolated intersection for the fixed-time control system using queueing system analysis. For this purpose, an appropriate definition of queuing system state is first provided, and the corresponding equilibrium equations are then extracted and solved, from which the limiting probabilities are derived. Finally, the performance criterion (average waiting time) for the queueing system is obtained, and the corresponding equation is considered as the objective function of the mathematical model.

TABLE 1. Literature review summary

	Intersection Type		**Decision Variables**				**Control System Type**		
	Isolated	**Arterial**	**Cycle Time**	**Green Time**	**Phases sequences**	**Red Time**	**Fixed Time**	**Actuated**	**Adoptive**
Webster [1]	*			*			*		
Chanloha [6]	*			*			*		
Akçelik [9]	*			*			*		
Pacheco [10]	*			*			*		
Yang [11]	*			*			*		
Zhang and Wang [17]	*	*		*			*		
Jiao et al. [18]	*		*	*				*	
Mirchandani and Zou [19] Chedjou and Kyamakya [20]	*		*	*					*
Ceylan and Bell [21]		*	*	*			*		
Wunderlich et al. [22] Wismans et al. [23]	*						*	*	
Ghavami et al. [24]	*			*	*		*	*	*
Ren et al. [25]	*			*			*		
Zhou and Cai [26]	*			*			*		
Dujardin et al [27]	*			*		*			*
Wu and Wang [28]	*			*			*		
Peñabaena et al. [29]		*	*	*			*		
Anusha et al. [30]	*		*	*			*		
Olszewski [31]	*		*	*			*		
Shirvani and Maleki [32]	*			*				*	
Lim et al. [33]	*			*			*		
Chin et al. [34]		*	*	*	*		*		
Current research	*		*	*			*		

The rest of the paper is organized as follows. Section 2 provides the problem definition and model assumptions. Section 3 presents equilibrium equations that are used to generate the mathematical model. Section 4 describes the results. Finally, section 5 summarizes the implications.

2. PROBLEM DEFINITION

Let us consider an isolated intersection with two straight movement phases at each of the two intersecting streets (phase 1 for north-to-south movements and phase 2 for east-to-west movements), where one line exists for vehicles to pass along in each phase. The arrivals in each phase are considered as a Poisson process with a rate of λ_i $(i = 1, 2)$, and the service is assumed to be an exponential process with a rate of μ_i $(i = 1, 2)$. Fixed-time timing is assumed for the traffic signal. In this strategy, the timing plan of traffic signals is adjusted to a preset time according to prior data. The decision variables are the green time and the red time. For analysis of the queueing system applied to the intersection, the green time (red time) for phase 1 is first assumed to have Erlang distribution with degrees of freedom g_1 (r_1) and rate A_1 (B_1). If the degrees of freedom tend to infinity in the limiting state, a constant will be obtained. Therefore, if the degrees of freedom for the green time (red time), *i.e.*, g_1 (r_1), tend to infinity, the green time (red time) will have a constant value with parameter A_1 (B_1), where A_1 (B_1) is the decision variable. The same analysis can be provided for phase 2; that is, we can assume in phase 2 that the red time (green time) has Erlang distribution with degrees of freedom g_2 (r_2) and rate A_2 (B_2). Again, if the degrees of freedom tend to infinity, a constant value with parameter A_2 (B_2) will be obtained for the red time (green time). The yellow time is assumed in the analysis to be constant for each phase and have Erlang distribution with degrees of freedom l and rate C, like the green and red times. Furthermore, if the degrees of freedom tend to infinity, a constant value with parameter C will be obtained for the yellow time. Unlike the green and red times, however, it is not a decision variable.

3. MATHEMATICAL MODEL

To obtain the equilibrium equations, the state of the system for each phase is defined as n_{i1}, n_{i2}, n_{i3}, where n_{i1} represents the number of vehicles in phase i $(0 \le n_{i1} \le N_i)$. Since it is not possible to solve the equilibrium equations for cases where system capacity is infinite, the intersection capacity is assumed to be a finite value N_i for $i = 1, 2$, providing a good approximation for the case of infinite capacity.

It should be noted that n_{i2} denotes the signal state for phase i, where $n_{i2} = 0$ represents the red signal, $n_{i2} = 1$ indicates the green light, and $n_{i2} = 2$ shows the yellow light. Moreover, n_{i3} represents the current state of the green or red light in phase i. In other words, when the traffic signal shows green light ($n_{12} = 1$) in phase 1, n_{13} increases stepwise from 1 to g_1, where $n_{13} = 1$ indicates the beginning of the green light, and $n_{13} = g_1$ denotes the end for phase 1. Once $n_{13} = g_1$, the signal for phase 1 shows yellow light ($n_{12} = 2$), and $n_{13} = 1$ again, increasing stepwise to 1. $n_{13} = l$ implies that the signal light in phase 1 should be changed from yellow to red and that in phase 2 should be changed from red to green. Therefore, $n_{12} = 0$ (phase 1 signal light is red), and $n_{22} = 1$ (phase 2 signal light is green). Then, n_{13} increases stepwise from 1 to r_1, and phase 1 signal light turns green when it reaches r_1. A similar analysis can be provided for phase 2. As a result, the bounds for n_{13} can be given as follows.

$$\text{Phase 1: } \begin{cases} 1 \le n_{13} \le r_1 \ , & n_{12} = 0 \\ 1 \le n_{13} \le g_1 \ , & n_{12} = 1 \\ 1 \le n_{13} \le l \ , & n_{12} = 2 \end{cases}$$

$$\text{Phase 2: } \begin{cases} 1 \le n_{23} \le g_2 \ , & n_{22} = 0 \\ 1 \le n_{23} \le r_2 \ , & n_{22} = 1 \\ 1 \le n_{23} \le l \ , & n_{22} = 2 \end{cases}$$

The system state is illustrated in the following diagrams for better understanding.

State (0, 1, 1): In this state (Figure 1), in phase 1, there is no vehicle, the light is green, and the first stage of the green time is dominant. The degrees of freedom for the green time is g_1, and n_{13} increases stepwise from 1 to g_1, when the light turns yellow. At that moment, $n_{13} = g_1$ indicates the end of the green time, $n_{12} = 2$, and $n_{13} = 1$. The latter denotes the beginning of the yellow time.

The equilibrium equation for state (0, 1, 1) is as follows.

$$\left(\lambda_1 + \frac{g_1}{A_1}\right)\pi_{(0,1,1)} = \frac{r_1}{B_1}\pi_{(0,0,r)} + \mu_1\pi_{(1,1,1)} \quad (1)$$

State $(n_{11}, 1, n_{13})$: In this state (Figure 2) in phase 1, there are n_1 vehicles. The signal indicates the green light because $n_2 = 1$, and the green signal is in mode n_3.

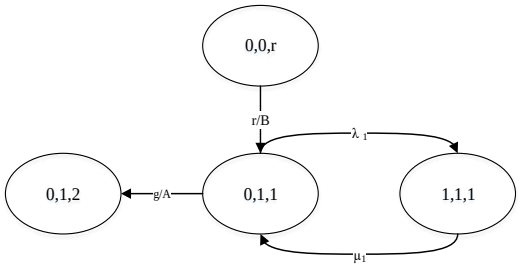


Figure 1. Diagram for state (0, 1, 1)

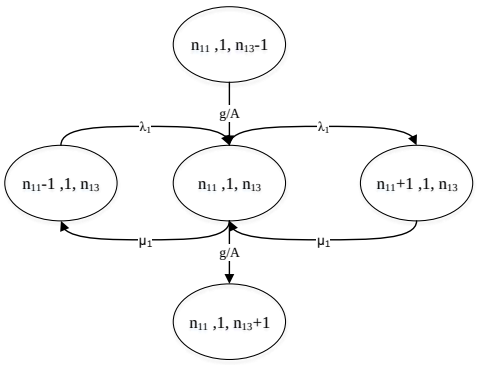


Figure 2. Diagram for state $(n_{11}, 1, n_{13})$

The equilibrium equation for state $(n_{11}, 1, n_{13})$ is as follows.

$$\left(\lambda_1 + \frac{g_1}{A_1} + \mu_1\right)\pi_{(n_{11},1,n_{13})} = \frac{g_1}{A_1}\pi_{(n_{11},1,n_{13}-1)} + \lambda_1\pi_{(n_{11}-1,1,n_{13})} + \mu_1\pi_{(n_{11}+1,1,n_{13})}, 1 \le n_{11} \le N_1, 2 \le n_{13} \le g \quad (2)$$

State (0, 0, 1): In this state (Figure 3) in phase 1, there is no vehicle, the signal shows red light since $n_2 = 0$, and the first stage of the red light is dominant. In this state and others where $n_2 = 0$, there is no service because the light is red.

The equilibrium equation for state (0, 0, 1) is as follows.

$$\left(\lambda_1 + \frac{r_1}{B_1}\right)\pi_{(0,0,1)} = \frac{l}{C}\pi_{(0,2,l)} \quad (3)$$

The equilibrium equation for state $(n_{11}, 0, n_{13})$ is as follows (Figure 4).

$$\left(\lambda_1 + \frac{r_1}{B_1}\right)\pi_{(n_{11},0,n_{13})} = \frac{r_1}{B_1}\pi_{(n_{11},0,n_{13}-1)} + \lambda_1\pi_{(n_{11}-1,0,n_{13})}, 1 \le n_{11} \le N_1, 2 \le n_{13} \le r \quad (4)$$

According to the above definitions, the equilibrium equations for phase 1 are as follows.

$$\left(\lambda_1 + \frac{r_1}{B_1}\right)\pi_{(0,0,1)} = \frac{l}{C}\pi_{(0,2,l)} \quad (5)$$

$$\left(\lambda_1 + \frac{r_1}{B_1}\right)\pi_{(0,0,n_{13})} = \frac{r_1}{B_1}\pi_{(0,0,n_{13}-1)}, 2 \le n_{13} \le r_1 \quad (6)$$

$$\left(\lambda_1 + \frac{r_1}{B_1}\right)\pi_{(n_{11},0,1)} = \frac{l}{C}\pi_{(n_{11},2,l)} + \lambda_1\pi_{(n_{11}-1,0,1)}, 1 \le n_{11} \le N_1 \quad (7)$$

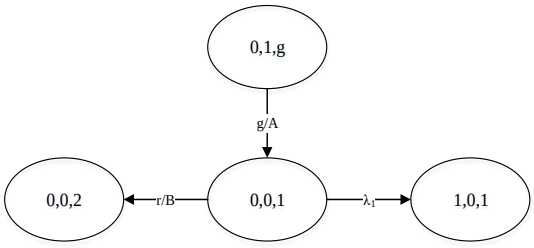


Figure 3. Diagram for state (0, 0, 1)

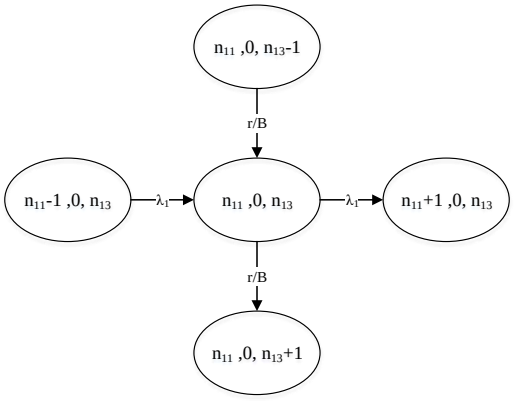


Figure 4. Diagram for state $(n_{11}, 0, n_{13})$

$$\left(\lambda_1 + \frac{r_1}{B_1}\right)\pi_{(n_{11},0,n_{13})} = \frac{r_1}{B_1}\pi_{(n_{11},0,n_{13}-1)} + \lambda_1\pi_{(n_{11}-1,0,n_{13})}, 1 \le n_{11} \le N_1, 2 \le n_{13} \le r_1 \quad (8)$$

$$\left(\lambda_1 + \frac{g_1}{A_1}\right)\pi_{(0,1,1)} = \frac{r_1}{B_1}\pi_{(0,0,r_1)} + \mu_1\pi_{(1,1,1)} \quad (9)$$

$$\left(\lambda_1 + \frac{g_1}{A_1}\right)\pi_{(0,1,n_{13})} = \frac{g_1}{A_1}\pi_{(0,1,n_{13}-1)} + \mu_1\pi_{(1,1,n_{13})}, 2 \le n_{13} \le g_1 \quad (10)$$

$$\left(\lambda_1 + \frac{g_1}{A_1} + \mu_1\right)\pi_{(n_{11},1,1)} = \frac{r_1}{B_1}\pi_{(n_{11},0,r_1)} + \lambda_1\pi_{(n_{11}-1,1,1)} + \mu_1\pi_{(n_{11}+1,1,1)}, 1 \le n_{11} \le N_1 \quad (11)$$

$$\left(\lambda_1 + \frac{g_1}{A_1} + \mu_1\right)\pi_{(n_{11},1,n_{13})} = \frac{g_1}{A_1}\pi_{(n_{11},1,n_{13}-1)} + \lambda_1\pi_{(n_{11}-1,1,n_{13})} + \mu_1\pi_{(n_{11}+1,1,n_{13})}, 1 \le n_{11} \le N_1, 2 \le n_{13} \le g_1 \quad (12)$$

$$\left(\lambda_1 + \frac{l}{C}\right)\pi_{(n_{11},2,n_{13})} = \frac{g_1}{A_1}\pi_{(n_{11},1,g_1)} + \lambda_1\pi_{(n_{11}-1,2,n_{13})} + \frac{l}{C}\pi_{(n_{11},2,n_{13}-1)}, 0 \le n_{11} \le N_1, 2 \le n_{13} \le l \quad (13)$$

$$\sum_{n_{13}=1}^{r_1}\sum_{n_{11}=0}^{N_1}\pi_{(n_{11},0,n_{13})} + \sum_{n_{13}=1}^{g_1}\sum_{n_{11}=0}^{N_1}\pi_{(n_{11},1,n_{13})} + \sum_{n_{13}=1}^{l}\sum_{n_{11}=0}^{N_1}\pi_{(n_{11},2,n_{13})} = 1 \quad (14)$$

The equilibrium equations for phase 2 can be presented along the same lines.

For example, consider state $(n_{21}, 2, n_{23})$, with n_1 vehicles in phase 2 (Figure 5). The signal indicates the yellow light because $n_2 = 2$, where the yellow signal is in mode n_3, and there is no service because the light is yellow.

The equilibrium equation for state $(n_{21}, 2, n_{23})$ would be as follows.

$$\left(\lambda_2 + \frac{l}{C}\right)\pi_{(n_{21},2,n_{23})} = \frac{r_2}{B_2}\pi_{(n_{21},1,g_2)} + \lambda_2\pi_{(n_{21}-1,2,n_{23})} + \frac{l}{C}\pi_{(n_{21},2,n_{23}-1)}, 0 \le n_{21} \le N_1 \quad 2 \le n_{23} \le l \quad (15)$$

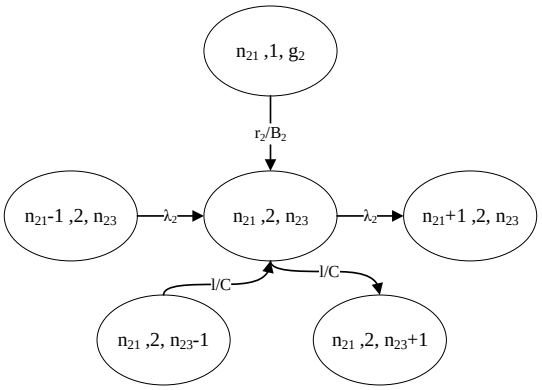


Figure 5. Diagram for state $(n_{21}, 2, n_{23})$

State $(n_{21}, 1, n_{23})$: In this state (Figure 6) in phase 2, there are n_1 vehicles. The signal indicates the green light because $n_2 = 1$, and the green signal is in mode n_3.

The equilibrium equation for state $(n_{21}, 1, n_{23})$ is as follows:

$$\left(\lambda_2 + \frac{r_2}{B_2} + \mu_2\right)\pi_{(n_{21},1,n_{23})} = \frac{r_2}{B_2}\pi_{(n_{21},1,n_{23}-1)} + \lambda_2\pi_{(n_{21}-1,1,n_{23})} + \mu_2\pi_{(n_{21}+1,1,n_{23})}, 1 \le n_{21} \le N_2, 2 \le n_{23} \le r_2 \quad (16)$$

The objective function of the model is to minimize the average waiting time at the intersection, considered as a weighted sum of the average waiting times in all the phases, as follows.

$$\min W = \frac{\lambda_1 w_1 + \lambda_2 w_2}{\lambda_1 + \lambda_2} \quad (17)$$

$$\lambda_1 w_1 = \sum_{n_{11}=0}^{N_1} n_{11}\pi_{(n_{11},n_{12},n_{13})}, \quad for \begin{cases} 1 \le n_{13} \le r_1, & n_{12} = 0 \\ 1 \le n_{13} \le g_1, & n_{12} = 1 \\ 1 \le n_{13} \le l, & n_{12} = 2 \end{cases} \quad (18)$$

$$\lambda_2 w_2 = \sum_{n_{21}=0}^{N_2} n_{21}\pi_{(n_{21},n_{22},n_{23})}, \quad for \begin{cases} 1 \le n_{23} \le g_2, & n_{22} = 0 \\ 1 \le n_{23} \le r_2, & n_{22} = 1 \\ 1 \le n_{23} \le l, & n_{22} = 2 \end{cases} \quad (19)$$

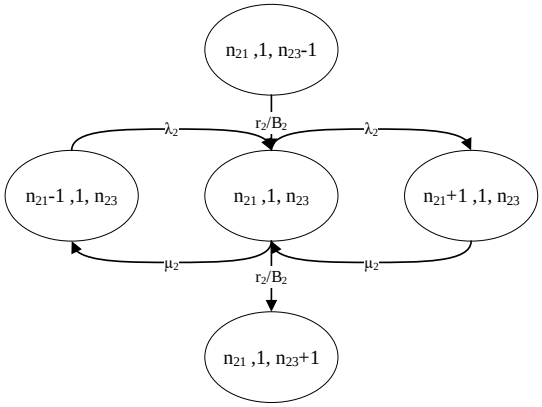


Figure 6. Diagram for state $(n_{21}, 1, n_{23})$

The set of constraints consists of all the equilibrium equations, each pertaining to a phase.
subject to:

$$\left(\lambda_1 + \frac{r_1}{B_1}\right)\pi_{(0,0,1)} = \frac{l}{C}\pi_{(0,2,l)} \quad (20)$$

$$\left(\lambda_1 + \frac{r_1}{B_1}\right)\pi_{(0,0,n_{13})} = \frac{r_1}{B_1}\pi_{(0,0,n_{13}-1)} \quad 2 \le n_{13} \le r_1 \quad (21)$$

$$\left(\lambda_1 + \frac{r_1}{B_1}\right)\pi_{(n_{11},0,1)} = \frac{l}{C}\pi_{(n_{11},2,l)} + \lambda_1\pi_{(n_{11}-1,0,1)}, 1 \le n_{11} \le N_1 \quad (22)$$

$$\left(\lambda_1 + \frac{r_1}{B_1}\right)\pi_{(n_{11},0,n_{13})} = \frac{r_1}{B_1}\pi_{(n_{11},0,n_{13}-1)} + \lambda_1\pi_{(n_{11}-1,0,n_{13})}, 1 \le n_{11} \le N_1, 2 \le n_{13} \le r_1 \quad (23)$$

$$\left(\lambda_1 + \frac{g_1}{A_1}\right)\pi_{(0,1,1)} = \frac{r_1}{B_1}\pi_{(0,0,r_1)} + \mu_1\pi_{(1,1,1)} \quad (24)$$

$$\left(\lambda_1 + \frac{g_1}{A_1}\right)\pi_{(0,1,n_{13})} = \frac{g_1}{A_1}\pi_{(0,1,n_{13}-1)} + \mu_1\pi_{(1,1,n_{13})}, \quad 2 \le n_{13} \le g_1 \quad (25)$$

$$\left(\lambda_1 + \frac{g_1}{A_1} + \mu_1\right)\pi_{(n_{11},1,1)} = \frac{r_1}{B_1}\pi_{(n_{11},0,r_1)} + \lambda_1\pi_{(n_{11}-1,1,1)} + \mu_1\pi_{(n_{11}+1,1,1)}, 1 \le n_{11} \le N_1 \quad (26)$$

$$\left(\lambda_1 + \frac{g_1}{A_1} + \mu_1\right)\pi_{(n_{11},1,n_{13})} = \frac{g_1}{A_1}\pi_{(n_{11},1,n_{13}-1)} + \lambda_1\pi_{(n_{11}-1,1,n_{13})} + \mu_1\pi_{(n_{11}+1,1,n_{13})}, 1 \le n_{11} \le N_1, 2 \le n_{13} \le g_1 \quad (27)$$

$$\left(\lambda_1 + \frac{l}{C}\right)\pi_{(n_{11},2,n_{13})} = \frac{g_1}{A_1}\pi_{(n_{11},1,g_1)} + \lambda_1\pi_{(n_{11}-1,2,n_{13})} + \frac{l}{C}\pi_{(n_{11},2,n_{13}-1)}, 0 \le n_{11} \le N_1, 2 \le n_{13} \le l \quad (28)$$

$$\sum_{n_{13}=1}^{r_1}\sum_{n_{11}=0}^{N_1}\pi_{(n_{11},0,n_{13})} + \sum_{n_{13}=1}^{g_1}\sum_{n_{11}=0}^{N_1}\pi_{(n_{11},1,n_{13})} + \sum_{n_{13}=1}^{l}\sum_{n_{11}=0}^{N_1}\pi_{(n_{11},2,n_{13})} = 1 \quad (29)$$

$$\left(\lambda_2 + \frac{g_2}{A_2}\right)\pi_{(0,0,1)} = \frac{l}{C}\pi_{(0,2,l)} \quad (30)$$

$$\left(\lambda_2 + \frac{g_2}{A_2}\right)\pi_{(0,0,n_{23})} = \frac{g_2}{A_2}\pi_{(0,0,n_{23}-1)}, 2 \le n_{23} \le g_2 \quad (31)$$

$$\left(\lambda_2 + \frac{g_2}{A_2}\right)\pi_{(n_{21},0,1)} = \frac{l}{C}\pi_{(n_{21},2,l)} + \lambda_2\pi_{(n_{21}-1,0,1)}, 1 \le n_{21} \le N_2 \quad (32)$$

$$\left(\lambda_2 + \frac{g_2}{A_2}\right)\pi_{(n_{21},0,n_{23})} = \frac{g_2}{A_2}\pi_{(n_{21},0,n_{23}-1)} + \lambda_2\pi_{(n_{21}-1,0,n_{23})}, 1 \le n_{21} \le N_2, 2 \le n_{23} \le g_2 \quad (33)$$

$$\left(\lambda_2 + \frac{r_2}{B_2}\right)\pi_{(0,1,1)} = \frac{g_2}{A_2}\pi_{(0,0,g_2)} + \mu_2\pi_{(1,1,1)} \quad (34)$$

$$\left(\lambda_2 + \frac{r_2}{B_2}\right)\pi_{(0,1,n_{23})} = \frac{r_2}{B_2}\pi_{(0,1,n_{23}-1)} + \mu_2\pi_{(1,1,n_{23})}, 2 \le n_{23} \le r_2 \quad (35)$$

$$\left(\lambda_2+\frac{r_2}{B_2}+\mu_2\right)\pi_{(n_{21},1,1)}=\frac{g_2}{A_2}\pi_{(n_{21},0,g_2)}+\lambda_2\pi_{(n_{21}-1,1,1)}+\mu_2\pi_{(n_{21}+1,1,1)},\ 1\le n_{21}\le N_2 \quad (36)$$

$$\left(\lambda_2+\frac{r_2}{B_2}+\mu_2\right)\pi_{(n_{21},1,n_{23})}=\frac{r_2}{B_2}\pi_{(n_{21},1,n_{23}-1)}+\lambda_2\pi_{(n_{21}-1,1,n_{23})}+\mu_2\pi_{(n_{21}+1,1,n_{23})},\ 1\le n_{21}\le N_2,\ 2\le n_{23}\le r_2 \quad (37)$$

$$\left(\lambda_2+\frac{l}{C}\right)\pi_{(n_{21},2,n_{23})}=\frac{r_2}{B_2}\pi_{(n_{21},1,g_2)}+\lambda_2\pi_{(n_{21}-1,2,n_{23})}+\frac{l}{C}\pi_{(n_{21},2,n_{23}-1)},\ 0\le n_{21}\le N_1\quad 2\le n_{23}\le l \quad (38)$$

$$\sum_{n_{23}=1}^{g_2}\sum_{n_{21}=0}^{N_2}\pi_{(n_{21},0,n_{23})}+\sum_{n_{23}=1}^{r_2}\sum_{n_{21}=0}^{N_2}\pi_{(n_{21},1,n_{23})}+\sum_{n_{23}=1}^{l}\sum_{n_{21}=0}^{N_2}\pi_{(n_{21},2,n_{23})}=1 \quad (39)$$

$$0\le\pi_{(n_{i1},n_{i2},n_{i3})}\le 1\quad,\ i=1,2 \quad (40)$$

$$A_{low}\le A\le A_{high} \quad (41)$$

$$B_{low}\le B\le B_{high} \quad (42)$$

Equations (18)-(27) show the equilibrium equations for phase 1, and Equations (28)-(37) show those for phase 2. The latter two determine the upper and lower bounds for the green and red times. If the green time in phase 1 is very long, the waiting time in phase 2 will increase substantially due to the long red time and vice versa. Therefore, the green and red times are required not to be longer than a specific amount.

4. CASE STUDY

In this section, in order to evaluate the queuing model, we compare its numerical results to those of simulating a single intersection. The simulation model for a fixed-time control system is coded in MATLAB. To solve the queuing model and calculate the optimal value of average waiting time at the intersection, we consider different values for the green and red times and obtain the average value of waiting time for each case by solving the equilibrium equations through MATLAB. The process continues until the optimal value is obtained for the average waiting time.

As mentioned earlier, the average waiting time at the intersection is obtained for different green and red times, where various values of red time are considered for each value of green time from 25 to 40 seconds (Values below the lower bound or above the upper bound would result in a significant increase in the average waiting time at the intersection). The average waiting time for each case of green and red time is calculated, and the red time associated with the minimum case is finally recorded in Table 1. For example, for the case where the green time is 34 seconds, average waiting time is calculated for different values of red time, and the optimal case is obtained, that is the minimum value of average waiting time at the intersection for the case in which the red time is 35 seconds. The optimal red times for the other green times are also obtained similarly. Moreover, the numerical solution of the queuing model is compared to the simulation results, as mentioned earlier, for assessment of the performance of the queueing model. The simulation model is coded in MATLAB and implemented ten times for each green or red signal length, each for 10,000,000 seconds.

The arrival and departure rates for each phase are as follows. This information concerns a two-phase intersection (Figure 7) in the city of Bojnurd, Iran.

$\lambda_1=0.25$, $\mu_1=0.67$ $\lambda_2=0.155$, $\mu_2=0.46$

The intersection capacity for each phase is assumed to be 50 cars (As stated earlier, this finite-capacity approach can provide a good approximation of the results for infinite capacity), and the degrees of freedom concerning Erlang distribution of the green, red, and yellow times are assumed to be 120. The yellow time is assumed to be 4 seconds for each phase (In the current state of the aforementioned intersection, the yellow time is 4 seconds).

Table 2 shows the green and red times and waiting time for phases 1 and 2 and the average waiting time at the intersection, obtained from the analytical model and simulation. The green and red times for phase 2 can be obtained easily from the results for phase 1, where the red time for phase 2 is obtained by adding the green and yellow times in phase 1. Moreover, the green time for phase 2 is obtained through subtraction of the yellow time for phase 2 from the red time for phase 1, Table 2.

$Green\ time_{phase\ 2}=Red\ time_{phase\ 1}-Yellow\ time_{phase\ 2}$

$Red\ time_{phase\ 2}=Green\ time_{phase\ 1}+Yellow\ time_{phase\ 1}$

Figure 8 compares the average waiting time obtained by the analytic model to that given by the simulation. As observed, there is an insignificant difference

Figure 7. Intersection under study

TABLE 2. Results of the model and simulation

Cycle	Phase 1			Phase 2			Average Waiting Time	
	Green Time	Red Time	Waiting Time	Green Time	Red Time	Waiting Time	Model	Simulation
1	25	27	30.9	23	29	32.7	31.6	31.5
2	26	28	29.8	24	30	33.4	31.2	31.1
3	27	29	25.8	25	31	38.9	30.8	30.8
4	28	30	25.3	26	32	39.1	30.6	30.5
5	29	31	25.8	27	33	37.9	30.4	30.4
6	30	31	24.2	27	34	40.2	30.3	30.2
7	31	32	24.4	28	35	39.5	30.2	30.1
8	32	33	26.8	29	36	35.4	30.1	30.0
9	33	34	27.0	30	37	35.0	30.1	30.0
10	34	35	25.3	31	38	37.7	30.1	30.0
11	35	35	25.3	31	39	37.8	30.1	30.1
12	36	36	24.3	32	40	39.6	30.2	30.1
13	37	37	24.6	33	41	39.3	30.2	30.1
14	38	38	27.6	34	42	34.7	30.3	30.2
15	39	39	26.6	35	43	36.7	30.4	30.4
16	40	40	26.7	36	44	36.8	30.6	30.5

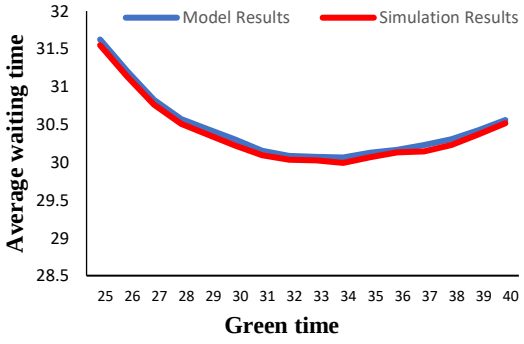


Figure 8. Average waiting time with respect to green time

between the results, and they are nearly matched. If the green time is longer or shorter than a certain amount, the average waiting time for the intersection will increase. This is because if the green time for phase 1 is long, the red time for phase 2 will increase, and so will the waiting time for phase 2 and the average waiting time for the whole intersection. Moreover, if the green time for phase 1 is short and that for phase 2 is long, the red time and the waiting time for phase 1 will increase. For this reason, the green and red times for each phase must be perfectly proportional to the arrival rate for that phase. At the intersection under study, the arrival rate for phase 1 is 1.6 times that for phase 2, and there should therefore not be much difference between the green times for the two phases. Figure 9 shows the waiting time for the two phases with respect to the cycle number. As can be seen, the waiting time for the former is always less than that for the latter because the arrival rate for phase 1 is greater than that for phase 2. Therefore, the green time for phase 1 is always greater than that for phase 2, leading to light traffic and a shorter average waiting time in phase 1 than in phase 2.

Table 3 summarizes the optimal solution obtained by the analytic model, along with a comparison of the average waiting time at the intersection in the optimal state to that in the current state. In the current state, the green time for phase 1 substantially differs from that for phase 2, while the arrival rates for the two phases do not differ significantly ($\lambda_1 = 0.25$, $\lambda_2 = 0.155$). A long red time and a short green time in phase 2 lead to heavy traffic and a dramatic increase in the vehicle waiting time in that phase and, consequently, in overall average waiting time at the intersection. The red times and green times for the two phases are set by the analytic model so that the average waiting times for the two phases and that for the entire intersection increase. However, in the current state, the average waiting time is very short in phase 1 and extremely long in phase 2, meaning that the traffic is light in phase 1 but heavy in phase 2.

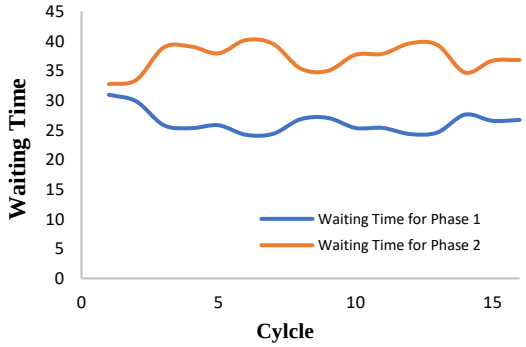


Figure 9. Waiting time for phases 1 & 2 with respect to cycle number

TABLE 3. A comparison of the optimal solution of the model and the current state

	Green Time Phase 1 (s)	Red Time Phase 1 (s)	Green Time Phase 2 (s)	Red Time Phase 2 (s)	Yellow Time (s)	Average Waiting Time (s)
Optimum result	34	35	31	38	4	30.1
Current state	50	30	26	54	4	99.2

From the Figures and tables above, the following results are implied.

1. The analytic model of intersection queue can significantly reduce the average waiting time at the isolated intersection. As indicated by the numerical results, the current conditions reduce average waiting time only in phase 1, but the analytic model sets the red times and green times for the two phases, so that average waiting time decreases in both phases, leading to relatively light traffic in them.
2. The analytic model can well estimate the average waiting time at the intersection, and its obtained results are very close to those given by the simulation of the fixed-time control system at the isolated intersection.

5. CONCLUSION

This paper has investigated the average waiting time obtained by a fixed-time signal control policy through analyses and simulations. We have analyzed a queuing model to estimate the average waiting time at a two-phase isolated intersection and calculated the optimal green and red times in a fixed-time traffic signal control system. The model has been solved for real traffic data collected from an isolated intersection in the city of Bojnurd, and the numerical results have been compared to those given by a simulation model. The comparison demonstrates that the analytical model approximates the simulation results very well. Moreover, it has been found that the analytical model obtained through analysis of the queuing model could substantially decrease the average waiting time at an isolated intersection. This paper has addressed the modeling of a fixed-time control system for a two-phase intersection. The method can be developed for adaptive or actuated control systems, and the results can be compared to those for the fixed-time control system. An assumption, and perhaps a drawback of the model, is that it considers a two-phase intersection, ignoring the turns taken to the left and right, which can greatly impact the average waiting time at the intersection. Furthermore, one can set green and red times for multiple intersections or a network instead of considering an isolated intersection, causing a decrease in the number of vehicle stops and average waiting time.

6. REFERENCES

1. Webster, F. V. "Traffic signal settings", Road research technical paper no. 39. Road Research Laboratory. (1958). https://trid.trb.org/view/113579
2. Miller, Alan J. "Settings for fixed-cycle traffic signals" ***Journal of the Operational Research Society*** 14.4, (1963), 373-386. https://doi.org/10.1057/jors.1963.61
3. Newell, Gordon Frank. "Approximation methods for queues with application to the fixed-cycle traffic light." ***Siam Review*** 7.2, (1965), 223-240. https://doi.org/10.1137/1007038
4. Heidemann, Dirk. "Queue length and delay distributions at traffic signals" ***Transportation Research Part B: Methodological,*** 28.5 (1994), 377-389. https://doi.org/10.1016/0191-2615(94)90036-1
5. Hu, X. N., L. C. Tang, and H. L. Ong. "AM/Dx/1 vacation queue model for a signalized intersection" ***Computers & Industrial Engineering*** 33.3-4 (1997), 801-804. https://doi.org/10.1016/S0360-8352(97)00240-4ctins B: Applications
6. Chanloha, , P., Usaha, W., Chinrungrueng, J. and Aswakul, C. "Performance comparison between queueing theoretical optimality and q-learning approach for intersection traffic signal control." 2012 Fourth International Conference on Computational Intelligence, Modelling and Simulation. IEEE, (2012) 172-177. https://doi.org/10.1109/CIMSim.2012.12
7. van den Broek, Mark S., et al. "Bounds and approximations for the fixed-cycle traffic-light queue." ***Transportation Science,*** Vol. 40, No. 4 (2006): 484-496. https://doi.org/10.1287/trsc.1050.0146
8. Habibi, M., A. Broumandnia, and A. Harounabadi. "Improvement of Multi-agent Routing Guidance with an Intelligent Traffic Light Scheduling and the Ability to Select Intermediate Destinations." ***International Journal of Engineering, Transactons A: Basics,*** Vol. 34, No. 4, (2021), 854-862. https://dx.doi.org/10.5829/ije.2021.34.04a.11
9. Akçelik, Rahmi, and Nagui M. Rouphail. "Overflow queues and delays with random and platooned arrivals at signalized intersections" ***Journal of Advanced Transportation*** 28.3 (1994), 227-251. https://doi.org/10.1002/atr.5670280305
10. Pacheco, António, Maria Lurdes Simões, and Paula Milheiro-Oliveira. "Queues with server vacations as a model for pretimed signalized urban traffic." ***Transportation Science*** 51.3 (2017), 841-851. https://doi.org/10.1287/trsc.2016.0727
11. Yang, Qiaoli, and Zhongke Shi. "The evolution process of queues at signalized intersections under batch arrivals." ***Physica A: Statistical Mechanics and its Applications*** Vol. 505, (2018), 413-425. https://doi.org/10.1016/j.physa.2018.03.059
12. Ghasemi, Jamal, and Jalil Rasekhi. "Traffic signal prediction using elman neural network and particle swarm optimization." ***International Journal of Engineering, Transactions B: Applications,*** Vol. 29, No. 11, (2016), 1558-1564. Doi: 10.5829/idosi.ije.2016.29.11b.09
13. Boon, M.A., Janssen, A.J.E.M., van Leeuwaarden, J.S. and Timmerman, R.W., "Pollaczek contour integrals for the fixed-cycle traffic-light queue." ***Queueing Systems*** Vol. 91, No. 1, (2019), 89-111. https://doi.org/10.1007/s11134-018-9595-9
14. Sumi, L., and V. Ranga. "Intelligent traffic management system for prioritizing emergency vehicles in a smart city." ***International Journal of Engineering, Transactions B: Applications,*** Vol. 31.2 (2018), 278-283. Doi:10.5829/ije.2018.31.02b.11
15. Amini, B., and J. Shahi. "The influence of urban network features on the quality of traffic service (research note)." ***International Journal of Engineering*** Vol. 11, No. 3, (1998), 167-174. http://www.ije.ir/article_71209.html
16. Faghri, A. "Signal design at isolated intersections using expert systems technology." ***International Journal of Engineering,*** Vol. 8, No. 4 (1995), 181-189. http://www.ije.ir/article_71132.html
17. Zhang, Guohui, and Yinhai Wang. "Optimizing minimum and maximum green time settings for traffic actuated control at isolated intersections." ***IEEE Transactions on Intelligent Transportation Systems*** Vol. 12, No. 1, (2010), 164-173. Doi: https://doi.org/10.1109/TITS.2010.2070795

18. Jiao, Pengpeng, Ruimin Li, and Zhihong Li. "Pareto front–based multi-objective real-time traffic signal control model for intersections using particle swarm optimization algorithm." ***Advances in Mechanical Engineering*** Vol. 8, No. 8, (2016), 1687814016666042. https://doi.org/10.1177%2F1687814016666042
19. Mirchandani, Pitu B., and Ning Zou. "Queuing models for analysis of traffic adaptive signal control." ***IEEE Transactions on Intelligent Transportation Systems*** Vol. 8, No. 1, (2007), 50-59. https://doi.org/10.1109/TITS.2006.888619
20. Chedjou, J.C. and Kyamakya, K., "A review of traffic light control systems and introduction of a control concept based on coupled nonlinear oscillators." ***Recent Advances in Nonlinear Dynamics and Synchronization,*** (2018), 113-149. Doi: 10.1007/978-3-319-58996-1_6
21. Ceylan, H. and Bell, M.G. "Traffic signal timing optimisation based on genetic algorithm approach, including drivers' routing." ***Transportation Research Part B: Methodological*** Vol. 38, No. 4, (2004), 329-342. https://doi.org/10.1016/S0191-2615(03)00015-8
22. Wunderlich, R., Liu, C., Elhanany, I. and Urbanik, T.,"A novel signal-scheduling algorithm with quality-of-service provisioning for an isolated intersection." ***IEEE Transactions on Intelligent Transportation Systems*** Vol. 9, No. 3, (2008), 536-547. https://doi.org/10.1109/TITS.2008.928266
23. Wismans, L., van Berkum, E. and Bliemer, M., "Dynamic traffic management measures to optimize air quality, climate, noise, traffic safety and congestion: effects of a single objective optimization." Transitions Towards Sustainable Mobility. Springer, Berlin, Heidelberg, 2011. 297-313. Doi: 10.1007/978-3-642-21192-8_16
24. Ghavami, Abouzar, Koushik Kar, and Satish Ukkusuri. "Delay analysis of signal control policies for an isolated intersection." 2012 15th International IEEE Conference on Intelligent Transportation Systems**.** IEEE, 2012. https://doi.org/10.1109/ITSC.2012.6338714
25. Ren, G., Huang, Z., Cheng, Y., Zhao, X. and Zhang, Y., "An integrated model for evacuation routing and traffic signal optimization with background demand uncertainty." ***Journal of Advanced Transportation,*** Vol. 47, No. 1, (2013), 4-27. https://onlinelibrary.wiley.com/doi/epdf/10.1002/atr.1211
26. Zhou, Zhanhong, and Ming Cai. "Intersection signal control multi-objective optimization based on genetic algorithm." ***Journal of Traffic and Transportation Engineering (English Edition)*** Vol. 1, No. 2, (2014), 153-158. https://doi.org/10.1016/S2095-7564(15)30100-8
27. Dujardin, Yann, Daniel Vanderpooten, and Florence Boillot. "A multi-objective interactive system for adaptive traffic control." ***European Journal of Operational Research*** Vol. 244, No. 2, (2015), 601-610. https://doi.org/10.1016/j.ejor.2015.01.059
28. Wu, Bin, and Dong Wang. "Traffic signal networks control optimize with PSO algorithm." 2016 12th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD). IEEE, 2016. https://doi.org/10.1109/FSKD.2016.7603179
29. Peñabaena-Niebles, Rita, Victor Cantillo, and José Luis Moura. "Impact of transition between signal timing plans in social cost based in delay, fuel consumption and air emissions." ***Transportation Research Part D: Transport and Environment*** 41, (2015), 445-456. https://doi.org/10.1016/j.trd.2015.10.018
30. Anusha, S. P., Sharma, A., Vanajakshi, L., Subramanian, S.C. and Rilett, L.R., "Model-based approach for queue and delay estimation at signalized intersections with erroneous automated data." ***Journal of Transportation Engineering*** Vol. 142, No. 5, (2016), 04016013. https://ascelibrary.org/doi/abs/10.1061/(ASCE)TE.1943-5436.0000835
31. Olszewski, Piotr S. "Modeling probability distribution of delay at signalized intersections." ***Journal of advanced transportation*** Vol. 28, No. 3, (1994), 253-274. https://doi.org/10.1002/atr.5670280306
32. Shiri, MJ Shirvani, and Hamid Reza Maleki. "Maximum green time settings for traffic-actuated signal control at isolated intersections using fuzzy logic*." **International Journal of Fuzzy Systems*** Vol. 19, No. 1, (2017), 247-256. https://doi.org/10.1007/s40815-016-0143-7
33. Lim, Dae Young, Xuan Zhou, and Kil To Chong. "The optimization of traffic signal control using LP and NLP." Applied Mechanics and Materials. Vol. 321. Trans Tech Publications Ltd, 2013. https://doi.org/10.4028/www.scientific.net/AMM.321-324.2241
34. Chin, Y.K., Yong, K.C., Bolong, N., Yang, S.S. and Teo, K.T.K., "Multiple intersections traffic signal timing optimization with genetic algorithm." 2011 IEEE International Conference on Control System, Computing and Engineering. IEEE, 2011. https://doi.org/10.1109/ICCSCE.2011.690569

Persian Abstract

چکیده

با افزایش جمعیت در شهرها، تعداد خودروهای موجود در جاده‌ها نیز افزایش می‌یابد که این منجر به کندی حرکت و تردد متراکم می‌شود. از جمله مشکلاتی که کم و بیش شهروندان شهرهای بزرگ با آن مواجه هستند، مسأله ترافیک در خیابان‌ها است. سیگنال‌های ترافیکی یکی از توانمندترین ابزارها برای کنترل ترافیک شهری هستند که نصب صحیح آنها می‌تواند هم جریان ترافیک و هم ایمنی کاربران را بهبود بخشد. برای کاهش اثر صف‌های طولانی خودروها در چهارراه‌ها تلاش زیادی صورت گرفته است. این تلاش‌ها عمدتاً در راستای بهینه‌سازی زمان‌بندی چراغ‌های راهنمایی بوده است. در این مقاله به کمک تحلیل سیستم صف، به محاسبه متوسط زمان انتظار در یک تقاطع منفرد و تنظیم زمان‌بندی طول چراغ سبز و قرمز پرداخته می‌شود. نوع سیستم کنترلی از نوع زمان ثابت و ورودی‌ها به تقاطع براساس فرآیند پواسان در نظر گرفته می‌شوند. پس از ارائه مدل ریاضی، داده‌های ترافیکی واقعی که از یک تقاطع دوفازه از شهر بجنورد تهیه شده است، برای حل مدل به کار گرفته می‌شوند و مقادیر بهینه طول چراغ سبز و قرمز بدست می‌آید. نتایج حاصل از شبیه‌سازی، قابلیت مدل پیشنهادی را در کنترل سیگنال‌های ترافیکی تایید می‌کند.

IJE TRANSACTIONS C: Aspects Vol. 34, No. 9, (September 2021) 2087-2096

International Journal of Engineering

Journal Homepage: www.ije.ir

Determining the Optimal Maintenance Strategy for Ammonium Hydroxide Production Unit Using Risk-Based Inspection and Analytic Hierarchy Process

R. Ghasemi[a], Y. Azimi*[a,b], Z. Ghasemi[c]

[a] *Department of Human Environment, College of Environment, Karaj, Iran*
[b] *Research Group of Environmental Engineering and Pollution Monitoring, Research Center for Environment and Sustainable Development, RCESD, Tehran, Iran*
[c] *Department of Industrial Engineering, Esfahan University of Technology, Esfahan, Iran*

PAPER INFO

Paper history:
Received 12 January 2021
Received in revised form 17 June 2021
Accepted 24 July 2021

Keywords:
Risk-based Inspection
Analytic Hierarchy Process
Maintenance
Ammonium Hydroxide

ABSTRACT

Ammonium Hydroxide production unit (AHPU) is one of the widely used and important units available in many industries such as oil and gas, petrochemical, and power plants. In this research, for the first time, a detailed systematic inspection and maintenance plan for AHPU is determined based on the prioritization of equipment using equipment's risk analysis. Equipment's failure risk is calculated based on probability and consequence of failure according to American Petroleum Institute Recommended Practice (API RP) 581. Different maintenance strategies were designed and compared considering several criteria like safety, cost, and feasibility using the analytic hierarchy process. Then risk-based optimal maintenance policy for each group of equipment was identified. Finally results showed that, the equipment identified as high or medium to high risk level such as liquid ammonia storage tank and its related pipes, should possesses preventive maintenance with an inspection period of 48 months. While for the medium risk equipment such as ammonium hydroxide storage tank and low-risk equipment such as ammonia absorption tank, should have corrective maintenance with inspection period of 72 and 90 months, respectively.

***doi**: 10.5829/ije.2021.34.09c.06*

1. INTRODUCTION

The occurrence of major accidents in industrial units leads to adverse effects on the main components of sustainable development, including the environment, economy, and society. According to recent studies, more than 80% of process unit accidents are related to 20% of equipment, which shows the importance of inspection in the industry [1-3]. Nowadays, there is an increasing attention to the risk-based methods for developing inspection, maintenance, and repair strategies in oil, gas, petrochemical, and power plant industries. The main reasons are the complexity of production processes in these industries, expensive equipment used, extra costs due to unexpected stops and overhauls, environmental contamination, and the importance of optimal usage of these processes [4, 5].

Risk-based inspection (RBI) is a management tool for identifying equipment risks and determining an inspection plan based on calculated risks. In RBI technique, first, the probability and consequence of equipment failure are calculated and the amount of risk is obtained by multiplying these two factors (probability and consequence). Then, based on the amount of the risk, equipment is prioritized and an inspection plan is defined for them. In this technique, unlike traditional inspection plans, which are based on a standard or manufacturer's suggestion, for each group of equipment, there will be a separate inspection schedule according to their level of risk [6].

In recent years, RBI and maintenance methods have become intertwined, and the tendency towards a risk-based approach for choosing an effective maintenance method has increased [1, 4, 7, 8]. On the other hand,

*Corresponding Author Email: yoosfazimi@gmail.com, azimi_yousef@rcesd.ac.ir (Y. Azimi)

Please cite this article as: R. Ghasemi, Y. Azimi, Z. Ghasemi, Determining the Optimal Maintenance Strategy for Ammonium Hydroxide Production Unit Using Risk-Based Inspection and Analytic Hierarchy Process, International Journal of Engineering, Transactions C: Aspects Vol. 34, No. 9, (2021) 2087-2096

choosing an appropriate maintenance plan is a very effective factor for decreasing maintenance time and cost in production units. A non-optimal maintenance plan may leads to unnecessary maintenance activities and wasting money and time or less maintenance activities resulting in more breakdowns and unwanted outages. Therefore, incorporating expert opinion and analysis in choosing the appropriate maintenance policy is an important factor that can improve the productivity of production units.

In 1990 Chen and Toyoda [9] developed a strategy for scheduling maintenance activities based on final risk. The RBI strategy was developed by the American Society of Mechanical Engineers (ASME) in 1991, [7]. In 2001 Dey [10] and Dey, Ogunlana and Naksuksakul [11] introduced a more general model for RBI and maintenance for offshore oil and gas pipelines. In these studies, the probability of failure in transmission pipelines is calculated by the AHP method by considering several criteria and sub-criteria among the known faults of pipelines. In 2007, Arunraj and Maiti [12] studied RBI methods and their application. The results of their study, which were performed on 25 RBI methods, showed that there is no unique method for conducting risk analysis and RBI. In 2009, Bertolini, Bevilacqua, Ciarapica and Giacchetta [13] reviewed the development of RBI and maintenance methods for oil refineries and proposed a method for RBI and maintenance. In 2010, Arunraj and Maiti [14] determined the appropriate maintenance plan for the benzene extraction unit of a chemical plant using the methods of Analytic Hierarchy Process (AHP) and goal programming. They provided a model for selecting risk-based maintenance (RBM) methods based on equipment failure and maintenance costs. This study illuminated that Condition-Based Maintenance (CBM) and Corrective Maintenance (CM) methods are suitable for high-risk and low-risk equipment, respectively. In 2012, Wang, Cheng, Hu and Wu [6] worked on the development of an RBM strategy using the FMEA method for the continuous catalytic reforming unit. The results of this study showed that the FMEA is a suitable approach for identifying critical equipment, and RBM policy can increase the reliability of high-risk equipment. In 2012, Kumar and Maiti [15] studied modeling the maintenance policy of an industrial unit using a fuzzy analysis network. The results show that the CBM is suitable when the risk of equipment is very high and the CM is preferred when the equipment risk is low and maintenance costs are significant. When both risk and maintenance costs are equally important, the Time-Based Maintenance (TBM) method is preferred. In 2018, Mohamed, Mahani and Razak [16] worked on the corrosion problem of equipment and lines of a floating platform in Petronance Company using the RBI method. Their research resulted in a 10-year and 15-year inspection and repair plan based on equipment risk classification. In 2019, Yazdi, Nedjati and Abbassi [17] developed a new maintenance strategy based on the intuitionistic Fuzzy AHP to obtain the weights of different investment factors for a separator system in an offshore process facility platform. The results indicate that the developed methodology estimates the risk more accurately, which enhances the reliability of future process operations. In 2020, Farizhendy, Noorzai and Golabchi [18] developed a model to select an optimal repair and maintenance method for jack-up drilling rigs in Iranian shipyards, they used genetic algorithm NSGA-II for solving multi-objective problem consists of cost, time, person-hour and environment-health. The results of this research help the experts of the Jack-up R&M make the right decision. In 2020, Eskandari, Charkhand and Gholami [1] provided a semi-quantitative management tool based on RBI on the de-ethanizer unit, to recognize and then prioritize the equipment's risks and, consequently, propose an inspection plan based on these risks. Their study results in saving a shutdown cost, inspection cost, reduction of failure, and increasing equipment reliability factor.

RBI is a regular method for determining and applying the failure rate of devices and equipment in the maintenance of facilities and making inspection decisions. RBI studies provide a more accurate understanding of the hazards and possible degradation mechanisms associated with pressurized reservoirs and piping. This information is very helpful in providing a complete asset integration plan and thus proper risk management. Moreover, RBI increases the time between technical inspections and periodic repairs. In this research, for the first time, an RBI plan will be determined for the equipment of the Ammonium Hydroxide Production Unit (AHPU), and the optimal maintenance strategy for the equipment of this production unit will be determined using the AHP method according to the risk level of equipment. In the AHPU, the consumed ammonia has a purity of 99.9%, and a 25% by weight solution of ammonia in water is prepared in this unit per day and at the same time, gaseous ammonia is sent to the consuming units. The main contributions of this study, which make this study original, are listed briefly in the following:

- Determining the risk types and levels governing equipment and process lines in AHPU for the first time
- Determining the intervals of technical inspection of equipment based on the level of risk of each equipment
- Determining maintenance policy based on the risk level of each equipment
- Increase the effectiveness of inspection operations and identify faults and their causes
- Determining the criteria affecting repair and maintenance in the AHPU

In the following, some related researches will be reviewed in section 2. In section 3, the proposed research method will be presented. In section 4, the results of implementing the proposed model in the AHPU will be presented and finally, the conclusion will be expressed in section 5.

2. THEORY OF MAINTENANCE, INSPECTION AND RISK ANALYSIS

Maintenance is a combination of management and engineering operations to maintain an object in a good working condition or to restore it to an acceptable condition. As stated, more than 80% of process unit accidents are related to problems in 20% of equipment (Figure). Common maintenance methods include Emergency Maintenance (EM), Breakdown Maintenance (BM), Preventive Maintenance (PM), CBM, TBM, Reliability-Based Maintenance (ReBM), and CM [4, 19]. One of the main objectives of each maintenance policy is to minimize hazards caused by unexpected equipment failure to humans and the environment.

Also, the maintenance policy should be cost-effective. Using a risk-based approach ensures that these goals are achieved. This approach uses the information obtained from the study of failures and their consequences [18]. Usually, risk can be determined qualitatively and quantitatively for the failure scenario. Based on the definition provided in American Petroleum Institute (API) 580 and API recommended practices 581(API RP 581) , risk can be calculated as follows [4, 7, 8]:

$$Risk = probability\ of\ failure \times Consequences\ of\ failure \quad (1)$$

After risk calculation for equipment, using the risk matrix, the priority of the assessed components is determined, and risk values are used to prioritize inspection and maintenance activities. In high and medium-risk areas, a centralized inspection, maintenance, and repair activity is required. That is while, in low-risk areas, maintenance-related activities are minimized to reduce the total inspection cost. RBI, which is a method for reducing the probability of an unpredictable failure, provides a set of suggestions on how to take preventive measures, including type, tools, and timing. The purpose of RBI is to identify more critical areas that need more attention in the inspection program. According to API RP 581, in qualitative analysis, first, the factor that indicates the possibility of a piece of equipment failure is identified, and then the factor for failure consequences is identified. By combining these two factors of failure and consequence, the position of the equipment in the risk matrix is determined and the equipment is ranked based on risk value [4, 8, 14].

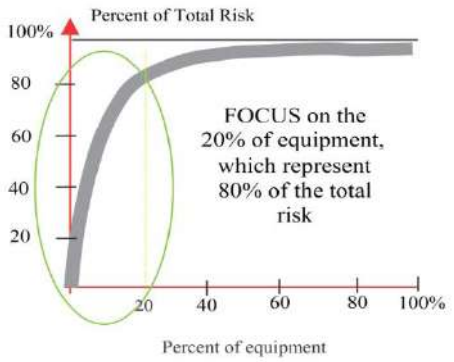


Figure 1. Total risk vs. quantity of equipment [4]

3. MATERIALS AND METHODOGY

3. 1. Ammonium Hydroxide Production Unit (AHPU) In this production unit, input ammonia has a purity of 99.9%. Liquid ammonia is discharged from the track to the storage tank then it is vaporized and reacted with water and converted to liquid ammonium hydroxide (AH). A 25% by weight solution of ammonia in water is prepared in this production unit, and at the same time, gaseous ammonia is sent to the consuming units. The schematic view of this production unit is shown in Figure 2. Liquid ammonia discharged from transportation tank to the storage tank. By passing through a vaporizer, it is converted to the gaseous form, then in the absorber ammonia gas is absorbed by water and liquid AH is produced. AH is widely used in various industries. The main application of this substance is in the production of chemical fertilizers, rubber, and plastics and as the main agent in household and industrial cleaners. It is also used in the chemical industry to neutralize acids. In the AHPU, there is a significant volume of ammonia gas and liquid and AH solution. The occurrence of failures in equipment and pipelines due to its toxic properties can greatly affect personnel health and the environment and waste of resources. AH is produced in almost the same method as in all AHPU; so it is necessary to ensure a reliable inspection and maintenance program. Identified damage mechanisms in the AHPU include; uniform corrosion, intergranular corrosion, galvanic corrosion, crevice corrosion, and pitting corrosion [20]. The equipment specifications of the AHPU are given in Table 1. The concept of the loop has been used to study the piping system.

3. 2. Methodology In Figure 3 steps designed for determining the optimal maintenance strategy for a typical AHPU are shown. Figure 3 was developed based on guidelines and procedures proposed in API RP 581 and intended to provide a systematic roadmap for

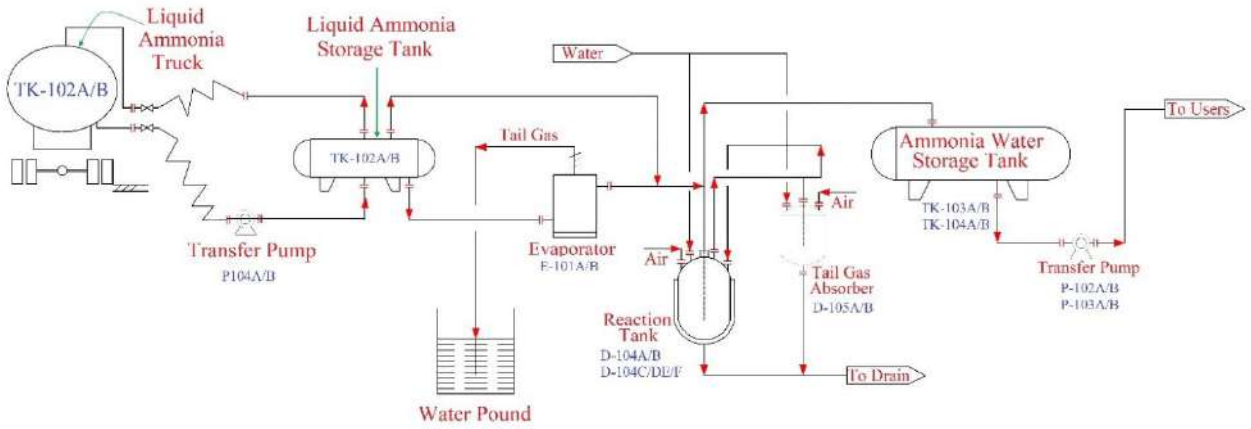


Figure 1. AHPU process flow diagram [21]

TABLE 1. Specifications of equipment and loops of AHPU [21]

Equipment ID	Type of Equipment	Operating pressure (barg)	Operating temperature (^{0}C)	Design Pressure (barg)	Design temperature (^{0}C)	Process fluid	Material of equipment
TK-102A/B	Pressure tank	18	45	22	85	Liquid ammonia	SA-516 Gr. 70N
E-101A/B/C	Evaporator	18	45	22	85	Gaseous and liquid ammonia	SA-516 Gr. 70N
D-104A/B/C/D/E/F	Double-walled tank under pressure	1	35	5	85	AH solution	SS-304
D-105A/B	Pressure tank	1	35	5	85	AH solution	SS-304
TK-103A/B	Pressure tank	1	35	5	85	AH solution	SS-304
TK-104A/B	Pressure tank	1	35	5	85	AH solution	SS-304
P-102A/B	Centrifugal pump	2.4	35	-	-	AH solution	SS-304
P-103A/B	Centrifugal pump	2.4	35	-	-	AH solution	SS-304
Loop 1	Pipe	18	45	-	-	Liquid ammonia	SS-304 300#
Loop 2	Pipe	18	45	-	-	Gaseous ammonia	SS-304 300#
Loop 3	Pipe	18	45	-	-	Liquid ammonia	SS-304 300#
Loop 4	Pipe	18	45	-	-	Gaseous ammonia	SS-304 300#
Loop 5	Pipe	1	45	-	-	Gaseous ammonia	SS-304 150#
Loop 6	Pipe	1	35	-	-	Gaseous ammonia	SS-304 150#
Loop 7	Pipe	1	35	-	-	AH solution	SS-304 150#
Loop 8	Pipe	1	35	-	-	AH solution	SS-304 150#
Loop 9	Pipe	2.4	35	-	-	AH solution	SS-304 150#

conducting similar researches. Research steps include data collection (technical and process information), identifying dominant corrosion mechanism of equipment, calculating probability and consequences of each equipment failure, equipment risk calculation, ranking and inspection plan determination and finally determining maintenance policy of equipment. In this research, according to API RP 581 standard, the equipment risk matrix has been determined and an inspection plan has been selected based on this matrix. Finally, using hierarchical analysis, the optimal method of maintenance is obtained. Based on the research topic and the prevailing conditions in the AHPU, the data collection method includes questionnaires, observations, interviews, available documents, information reports, and databases. According to Figure 3, in this study, first, the technical data of the equipment and process information of the AHPU are gathered.

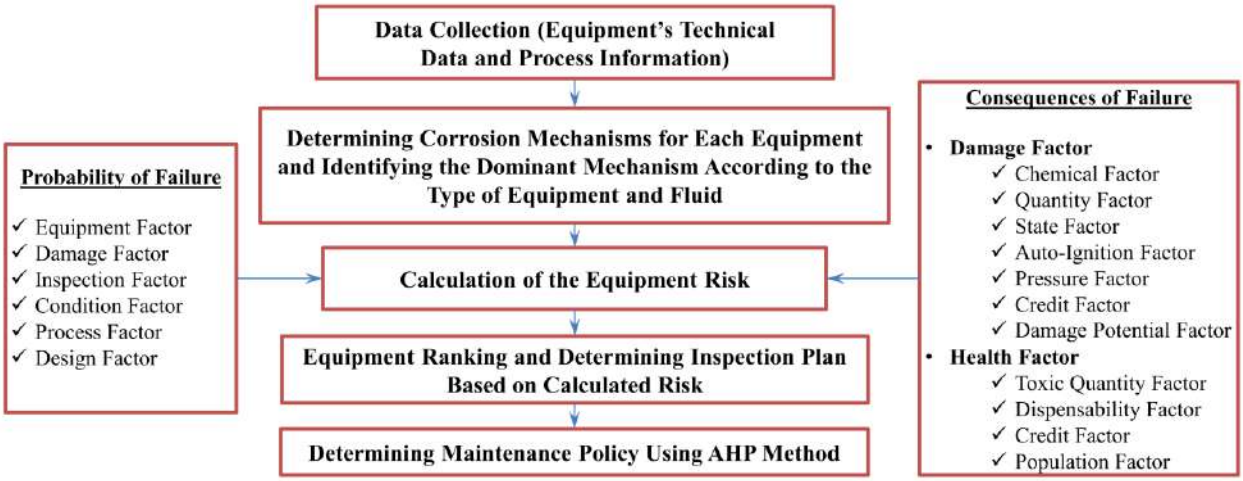


Figure 3. Risk-based inspection and maintenance planning process [21]

This data includes temperature, pressure, and type of equipment, equipment material, equipment safety systems, design-related data, equipment inspection history, and fluid properties, including corrosion, flammability, and toxicity. Then according to the type of equipment and the type of fluid carried with in it, the dominant corrosion mechanism for each equipment is determined. The risk of each equipment is calculated and the equipment is ranked based on the resulting risk. Then, according to the level of risk for each of the equipment and the results of previous inspections, the appropriate inspection plan is determined. Finally, the maintenance policy is determined using the AHP method.

3. 2. 1. Calculating Equipment Risk In this study, a qualitative risk analysis is performed according to the method given in API RP 581 with the help of a series of guideline tables designed to evaluate an industrial unit in a short time. The basics of qualitative risk analysis are the same as the semi-quantitative method, but this method requires less detail for risk analysis, so with less accuracy in comparison with the quantitative method. But it can be used to prioritize the inspection program. In the qualitative risk analysis, the risk level for each equipment is obtained from the classification of failure probability and the maximum categorization of health and injury outcomes in the risk matrix. These levels are divided into 4 levels including Low Risk, Medium Risk, Medium to High Risk, and High Risk. In the used risk matrix, as shown in Table 2 the failure consequence is displayed on the horizontal axis and the probability category of failure is displayed on the vertical axis. Then the position of each equipment on this matrix is determined and the risk level of the equipment is identified accordingly [8].

TABLE 2. Risk matrix [8]

Failure probability category						
	5	**Moderate High (MH)**	**Moderate High (MH)**	**Moderate High (MH)**	**High (H)**	**High (H)**
	4	Moderate (M)	Moderate (M)	Moderate High (MH)	Moderate High (MH)	High (H)
	3	Low (L)	Low(L)	Moderate (M)	Moderate High (MH)	High (H)
	2	Low (L)	Low (L)	Moderate (M)	Moderate (M)	Moderate High (MH)
	1	Low (L)	Low (L)	Moderate (M)	Moderate (M)	Moderate High (MH)
		A	**B**	**C**	**D**	**E**
		Consequences of Failure				

3. 2. 1. 1. Determining The Probability of Failure The method of calculating the probability of equipment failure using six factors shown in Figure 2 is stated in section A of the API RP 581 standard. These six factors, which affect the probability of the occurrence of a large leakage include the Equipment Factor (EF), DF, Inspection Factor (IF), Condition Factor (CoF), Process Factor (PF), and Mechanical Design Factor (MDF). The value of each factor is determined based on the guideline tables presented in API RP 581 section A. The sum of these six factors indicates the probability of total failure (Equation (2)). After calculating the failure probability, its category is determined for each equipment using Table 3 [8].

$$Probability\ of\ failure = EF+DF+IF+CoF+PF+MDF \quad (2)$$

TABLE 3. Determining failure probability categories [8]

Probability of Failure	Category of Probability of Failure
0-15	1
16-25	2
26-35	3
36-50	4
51-75	5

3. 2. 1. 2. Determining The Consequences of Failure Determining the consequences of failure includes the calculation of two separate factors namely, damage factor and health factor (the toxicity consequences of the fluids). To determine the consequence, seven sub-factors are determined based on guideline tables presented in API 581 sections B, including Chemical Factor (ChF), Quantity Factor (QF), State Factor (SF), Auto-Ignition Factor (AF), Pressure Factor (PRF), Credit Factor (CRF), and Damage Potential Factor (DPF). The health factor value obtained by determining the four sub-factors named Toxic Quantity Factor (TQF), Dispensability Factor (DIF), CRF, and Population Factor (PPF) using guideline tables presented in API 581 sections C. These two factors are calculated for each chemical using Equations (3) and (4). According to Table 4 the factor with the higher category will be used as the consequences of failure in risk calculation in Equation (1). In Table 4 the category of the consequence of failure is indicated alphabetically in the way that "A" is the lowest and "E" is the highest one. Bringing the definition of each of the factors in the two preceding sections is out of the scopes of this paper and can be fined as follows [7, 8]:

$$Damage\ Factor = ChF + QF + SF + AF + PRF + CRF + DPF \quad (3)$$

$$Health\ Factor = TQF + DIF + CRF + PPF \quad (4)$$

3. 2. 2. Determining the Inspection Period Usually, based on the equipment risk level and type of equipment and its failure mechanisms, the inspection programs are developed for periodic inspections, including in-service inspections and inspections during overhaul. In Table 5 the relationship between risk level, inspection level, and recommended equipment inspection period is given. Based on this table and according to the selected level of inspection and obtained risk level for equipment, the inspection period for each equipment can be determined [22].

TABLE 2. Determining the consequence of failure [7, 8]

Health factor	Damage factor	Category of Consequences of failure
>10	0-19	A
10-19	20-34	B
20-29	35-49	C
30-39	50-79	D
<40	<79	E

3. 2. 3. Determining the Inspection Period Choosing the most appropriate maintenance policy is a complicated task due to the diversity of possible alternatives and multiple criteria. For selecting the best maintenance strategy, the risk level obtained through RBI can be used to identify the equipment in critical condition and, if necessary, more maintenance resources can be allocated to them. The next step is to decide on the maintenance policy from various possible approaches according to the risk level of equipment. For this purpose, the AHP technique can be utilized [23]. The AHP is a structured method for handling complex decisions. AHP was developed by Thomas L. Saaty in the 1970s. The AHP gives a complete and rational agenda for structuring a decision problem, representing and quantifying its elements, relating those elements to overall goals, and evaluating alternative solutions. AHP has three main levels. For maintenance problems; the first level is the goal (in this case, choosing the best maintenance policy for each mentioned risk level). The second level consists of decision criteria and the third level is alternatives (in this case, applicable maintenance policies in AHPU).

Performing AHP analysis consists of three stages namely: building decision models' structure; carrying out comparative evaluation on the alternatives and the criteria; the combination of the priorities. In the first

TABLE 3. Relationship between risk level, inspection level and recommended inspection period [22]

			Risk Level			
			L	M	MH	H
Inspection Level	1	-There was no previous inspection history at all -There was not enough inspection history	60	48	36	24
	2	-At least 1 inspection has been performed -No depreciation found -Has at least 2 inspection records	90	72	48	36
	3	-There is at least 1 inspection history based on RBI -Unforeseen depreciation not found -There is at least 3 inspection records	120	96	72	48
	4	-At least 2 inspections have been performed based on RBI -Only predicted depreciation has occurred	120	120	96	60

stage, the decision problem is built as a hierarchy, by breaking down the complex decision construction problem into the hierarchy of objectives at the top, in the middle, and the alternatives at the bottom. In the second stage, the pairwise comparisons of the alternatives and criteria are carried out. Let $C=(C_j \mid j=1,2,...,n)$ be the set of criteria, an $(n\times n)$ evaluation matrix A, with elements of a_{ij} $(i,j=1,2,...,n)$ showing the quotient of weights of the criteria, summarizes the pairwise comparison on n criteria, as depicted in Equation (5). In this study, the pairwise comparison mechanism for each criterion and alternative is done based on Table 6 which was proposed by Saaty.

$$A=\begin{bmatrix} a_{11} & . & a_{1n} \\ . & . & . \\ a_{n1} & . & a_{nn} \end{bmatrix}, a_{ii}=1, a_{ji}=\frac{1}{a_{ij}}, a_{ij}\neq 0 \quad (5)$$

Finally, the normalized relative weight for each criterion is obtained by the right eigenvector (ω) corresponding to the largest eigenvalue (λ_{max}), as follows (Equation (6))

$$A\omega=\lambda_{max}\omega \quad (6)$$

In this study to select the criteria affecting the maintenance policy, 13 different policies were considered and a questionnaire was prepared to get experts' opinions on the importance of each criterion. These 13 criteria include surroundings and buildings safety, equipment failure number or intervals, ease of access of the inspection team to the equipment, toxic substances dispersion effects, added value, production stoppage time, accessibility of the equipment in risky areas, spare parts cost, expert manpower cost, equipment safety, equipment importance in the processes, feasibility, and availability of equipment, and personnel safety. The four criteria with the highest average importance namely; safety, cost, added value, feasibility, and availability of equipment were selected for choosing a maintenance policy. Accordingly, for maintenance alternatives, PM, CM, ReBM, and CBM were considered. After determining the criteria and options of the hierarchical model, the proposed model is depicted in Figure 4.

4. RESULTS AND DISCUSSION

In this study, 30 personnel working in the AHPU were identified as a statistical population based on the type of specialization and education for data gathering. They were specialists in various fields of maintenance, inspection, safety, engineering, and production. The required information was collected through questionnaires, interviews, and observations.

TABLE 6. Methodology for Judgment in AHP [23]

Judgment	Equally	Moderately	Strongly	Very strongly	Extremely
Score	1	2,3	4,5	6,7	8,9

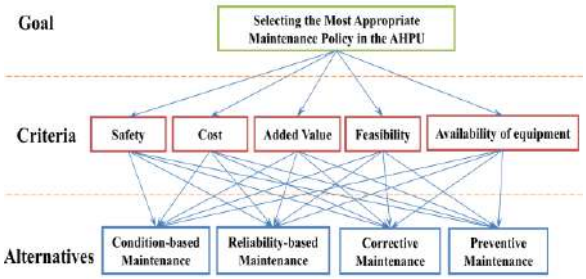


Figure 4. The proposed AHP model

4. 1. Determining Failure Probability Category and Consequence of Failure To determine the number of factors affecting the equipment failure, three experts were surveyed, and with the summation of amounts for included factors, the amount of failure probability and probability category for each equipment was determined according to Table 3. Due to the low flammability of ammonia, the damage factor is ignored in calculating the consequence of failure. Similarly, the values of health factors for equipment have been determined through interviews with relevant experts, and the failure consequence has been determined for each equipment based on Table 4. The Results of failure probability and consequence of failure category are shown in Table 7.

4. 2. Risk Level and Inspection Interval After identifying the failure probability category and consequence of failure for equipment, the risk level of each equipment is determined according to Table 3. On the other hand, since during the inspection, no damage was observed in the equipment and piping system of the AHPU, inspection level 2 for the equipment and piping system is considered, and inspection interval is identified according to Table 5, as shown in Table 8.

4. 3. Selecting the Optimal Maintenance Strategies using AHP Method As mentioned in the previous sections, to implement the AHP method, the criteria's weight must be determined according to pairwise comparisons. In this study, the expert choice software was used that works according to the AHP.

4. 3. 1. Pairwise Comparison Matrixes Calculation As mentioned before, risk matrix output consists of four main risk rating scales. A suitable maintenance policy must be assigned to each risk rating scale by calculating each policy priority using pairwise comparison matrixes. The most important point is that the ranking of criteria and alternatives are different for each risk rating scale, so

the pairwise comparison matrix must be calculated in each risk rating scale. As mentioned before, rankings are assigned to each criterion and alternative based on Saaty ranking table (Table 6). For the AHPU, the weights obtained for the selected criteria for each risk level are as described in Table 9. Accordingly, pairwise comparisons of maintenance strategies based on criteria chosen for different risk level has performed. Table 10 shows these pairwise comparisons in the low-risk range.

4. 3. 1. Pairwise Comparison Matrixes Calculation

As mentioned before, risk matrix output consists of four main risk rating scales. A suitable maintenance policy each policy priority using pairwise comparison matrixes. The most important point is that the ranking of criteria and alternatives are different for each risk rating scale, so the pairwise comparison matrix must be calculated in each risk rating scale. As mentioned before, rankings are assigned to each criterion and alternative based on Saaty ranking table (Table 6). For the AHPU, the weights obtained for the selected criteria for each risk level are as

TABLE 4. Classification of equipment probability of failure and consequences of failure in the studied AHPU

Equipment ID	LF	MDF	PF	CCF	IF	DF	EF	LC	TQF	DIF	CFR	PPF	HCF	HCC	DCC	OCC
TK-102A/B	23	0	3	6	-4	18	0	2	27	0	-0.5	20	47	E	-	E
E-101A/B/C	21	0	3	6	-4	16	0	2	15	0	0	20	35	D	-	D
D-104A/B/C/D/E/F	19	0	4	6	-4	13	0	2	15	-5	-1	7	16	B	-	B
D-105A/B	17	0	2	6	-4	13	0	2	15	-5	-1	7	16	B	-	B
TK-103A/B	20	0	2	6	-4	16	0	2	20	-5	-1	7	21	C	-	C
TK-104A/B	20	0	2	6	-4	16	0	2	20	-5	-1	7	21	C	-	C
P-102A/B	28	0	2	6	0	20	0	3	20	-5	-1.25	7	21	C	-	C
P-103A/B	28	0	2	6	0	20	0	3	20	-5	-1.25	7	21	C	-	C
Loop 1	28	0	3	9	0	16	0	3	15	0	0	20	35	D	-	D
Loop 2	28	0	3	9	0	16	0	3	15	0	0	20	35	D	-	D
Loop 3	28	0	3	9	0	16	0	3	15	0	0	20	35	D	-	D
Loop 4	28	0	3	9	0	16	0	3	15	0	0	20	35	D	-	D
Loop 5	28	0	3	9	0	16	0	3	15	0	-1.5	20	35	D	-	D
Loop 6	27	0	2	9	0	16	0	3	15	0	-1.5	7	21	C	-	C
Loop 7	27	0	2	9	0	16	0	3	15	-5	-1.5	7	16	B	-	B
Loop 8	27	0	2	9	0	16	0	3	20	-5	-1.25	7	21	C	-	C
Loop 9	3	27	0	2	9	0	16	0	20	-5	-1.25	7	21	C	-	C

TABLE 8. Recommended inspection interval for equipment of AHPU

Equipment title	Probability category	Consequence of failure	Risk Level	Inspection interval (month)
Under pressure tank	2	E	MH	48
Evaporator	2	D	M	72
Double-walled under pressure tank	2	B	L	90
Centrifugal pump	3	C	M	72
Piping line 1 to 4	3	D	MH	48
Piping line 6	3	C	M	72
Piping line 7	3	B	L	90
Piping lines 8 and 9	3	C	M	72

TABLE 9. The weight of each criterion for different risk levels

Criteria	Low Risk	Moderate Risk	Moderate to High Risk	High Risk
Safety	0.16	0.17	0.29	0.35
Added value	0.22	0.28	0.22	0.21
Feasibility	0.14	0.14	0.19	0.18
Cost	0.28	0.22	0.16	0.15
Availability of equipment	0.2	0.19	0.14	0.11

described in Table 9. Accordingly, pairwise comparisons of maintenance strategies based on criteria chosen for different risk level has performed. Table 10 shows these pairwise comparisons in the low-risk range.

According to values obtained in Tables 9 and 10, the priority of each maintenance strategy was calculated in the low-risk range, which is illustrated in Table 11. As can be seen, for low-risk equipment, the best strategy is the CM. Similarly, pairwise comparisons of maintenance strategies for other risk levels were performed and the relative importance of each strategy was determined, which is shown in Table 12.

According to the obtained results in section 4, the appropriate inspection levels for equipment with low risk, medium risk, medium to high-risk levels are 90, 72, and 48 months, respectively. It was also found that the optimal maintenance policies for equipment with different risk levels are shown in Table 13.

TABLE 10. Comparisons of maintenance alternatives in the low-risk range

	Criteria				
Maintenance Alternatives	**Safety**	**Feasibility**	**Added value**	**Cost**	**Availability of equipment**
PM	0.216	0.315	0.290	0.248	0.23
CM	0.281	0.228	0.220	0.345	0.3
RBM	0.269	0.238	0.262	0.224	0.17
CBM	0.234	0.219	0.290	0.183	0.3

TABLE 11. Optimal maintenance strategy in the low-risk range

Maintenance Alternatives	**The relative importance of each strategy**
PM	0.271
CM	0.296
RBM	0.233
CBM	0.200

TABLE 12. Optimal maintenance strategy based on risk level

Maintenance Alternatives	**The relative importance of each strategy for different risk levels**		
	High Risk	**Moderate to High Risk**	**Moderate Risk**
PM	0.298	0.300	0.286
CM	0.193	0.193	0.302
RBM	0.279	0.277	0.222
CBM	0.229	0.230	0.189

TABLE 13. The most appropriate maintenance policy based on risk levels

Risk Level	**Maintenance Strategy**
Low Risk	CM
Moderate Risk	CM
Moderate to High Risk	PM
High Risk	PM

5. CONCLUSION

In this research, an appropriate equipment inspection and maintenance plan for equipment in a typical AHPU was determined based on RBI and a multi-criteria decision system. AH is a widely used chemical in neutralizing acids in different industries producing chemical fertilizers, rubbers, plastics and exists as the main agent in household and industrial cleaners. AHPU is usually a separate part of different chemical industries and due to its important role in the chemical processes needs special inspection and maintenance strategy. Hence, contrary to traditional methods, which consider a fixed period for equipment inspection mostly based on the manufacturers' suggestions, in this study, the knowledge of the experts and equipment working and environmental conditions were used to determine the optimum period of inspections in a typical AHPU. The procedure considered in API 581 was employed to categorize the equipment based on their failure risk. Then based on RBI concepts, experts' opinions and using AHP decision methodology the appropriate maintenance policy considering were determined for each equipment categorized based on failure risk. Hence the equipment with high or medium to high risk level, such as liquid ammonia storage tank and the pipes entering or leaving it, needs PM plan with inspection period of 48 months. The equipment such as AH storage tank which has medium risk level needs CM plan with inspection period of 72 months. The equipment with low risk, such as ammonia absorption tank, requires CM plan with inspection period of 90 months. It is expected that by employing the proposed maintenance plan, the safety of the process in AHPU will be enhanced and additional cost imposed by equipment breakdowns will be reduced. Due to the widespread use of gaseous ammonia and AH solution in the industry; the obtained results can be use as guild for other similar production units as well

6. REFERENCES

1. Eskandari, D., Charkhand, H. and Gholami, A., "A semi-quantitative approach development for risk-based inspection in a petrochemical plant", ***Open Access Macedonian Journal of Medical Sciences***, Vol. 8, No. E, (2020), 425-433, doi: 10.3889/oamjms.2020.4391

2. Koteleva, N., Buslaevb, G., Valneva, V. and Kunshinb, A., "Augmented reality system and maintenance of oil pumps", ***International Journal of Engineering***, Vol. 33, No. 8, (2020), doi: 10.5829/ije.2020.33.08b.20.
3. Drożyner, P., "The impact of the implementation of management system on the perception of role and tasks of maintenance services and effectiveness of their functioning", ***Journal of Quality in Maintenance Engineering***, (2020), doi: 10.1108/JQME-09-2019-0089.
4. Calixto, E., "Gas and oil reliability engineering: Modeling and analysis, Gulf Professional Publishing, (2016).
5. Rahimi Komijani, H., Shahin, M. and Jabbarzadeh, A., "Optimal policy of condition-based maintenance considering probabilistic logistic times and the environmental contamination issues", ***International Journal of Engineering***, Vol. 31, No. 2, (2018), 357-364, doi: 10.5829/ije.2018.31.02b.21.
6. Wang, Y., Cheng, G., Hu, H. and Wu, W., "Development of a risk-based maintenance strategy using fmea for a continuous catalytic reforming plant", ***Journal of Loss Prevention in the Process Industries***, Vol. 25, No. 6, (2012), 958-965, doi: 10.1016/j.jlp.2012.05.009.
7. Inspection, R.-B., "Api recommended practice 580", ***American Petroleum Institute***, Vol., No., (2009), doi.
8. Segment, A.P.I.D., "Risk-based inspection technology, American Petroleum Institute, (2008).
9. Chen, L. and Toyoda, J., "Maintenance scheduling based on two level hierarchical structure to equalize incremental risk", in Conference Papers Power Industry Computer Application Conference, IEEE. (1989), 431-437.
10. Dey, P., "A risk based model for inspection and maintenance of cross country petroleum pipeline", ***Journal of Quality in Maintenance Engineering***, (2001), doi: 10.1108/13552510110386874.
11. Dey, P.K., Ogunlana, S.O. and Naksuksakul, S., "Risk based maintenance model for offshore oil and gas pipelines: A case study", ***Journal of Quality in Maintenance Engineering***, (2004), doi: 10.1108/13552510410553226.
12. Arunraj, N. and Maiti, J., "Risk-based maintenance—techniques and applications", ***Journal of hazardous materials***, Vol. 142, No. 3, (2007), 653-661, doi: 10.1016/j.jhazmat.2006.06.069.
13. Bertolini, M., Bevilacqua, M., Ciarapica, F.E. and Giacchetta, G., "Development of risk-based inspection and maintenance procedures for an oil refinery", ***Journal of Loss Prevention in the Process Industries***, Vol. 22, No. 2, (2009), 244-253, doi: 10.1016/j.jlp.2009.01.003.
14. Arunraj, N. and Maiti, J., "Risk-based maintenance policy selection using ahp and goal programming", ***Safety Science***, Vol. 48, No. 2, (2010), 238-247, doi: 10.1016/j.ssci.2009.09.005.
15. Kumar, G. and Maiti, J., "Modeling risk based maintenance using fuzzy analytic network process", ***Expert Systems with Applications***, Vol. 39, No. 11, (2012), 9946-9954, doi: 10.1016/j.eswa.2012.01.004.
16. Mohamed, N.A., Mahani, Z. and Razak, L.A.A., "Approach to corrosion management program and risk based inspection development for the world's first floating lng", in Offshore Technology Conference Asia, Offshore Technology Conference. (2018).
17. Yazdi, M., Nedjati, A. and Abbassi, R., "Fuzzy dynamic risk-based maintenance investment optimization for offshore process facilities", ***Journal of Loss Prevention in the Process Industries***, Vol. 57, (2019), 194-207, doi: 10.1016/j.jlp.2018.11.014.
18. Farizhendy, M.M., Noorzai, E. and Golabchi, M., "Implementing the nsga-ii genetic algorithm to select the optimal repair and maintenance method of jack-up drilling rigs in iranian shipyards", ***Ocean Engineering***, Vol. 211, (2020), 107548, doi: 10.1016/j.oceaneng.2020.107548.
19. Mokhtari, H., Molla-Alizadeh, S. and Noroozi, A., "A reliability based modelling and optimization of an integrated production and preventive maintenance activities in flowshop scheduling problem", ***International Journal of Engineering***, Vol. 28, No. 12, (2015), 1774-1781, doi: 10.5829/idosi.ije.2015.28.12c.10.
20. Cramer, S.D., Covino, B.S. and Moosbrugger, C., "Asm handbook volume 13b: Corrosion: Materials, ASM International Materials Park, OH, Vol. 13, (2005).
21. Ghasemi Vincheh, R., "Maintenance strategy selection at an ammonium hydroxide unit using a risk based inspection approach in pursuit of sustainable development", Department of Human Environment, College of Environment, Karaj, Iran, MSc, (2020),
22. Shuai, J., Han, K. and Xu, X., "Risk-based inspection for large-scale crude oil tanks", ***Journal of Loss Prevention in the Process Industries***, Vol. 25, No. 1, (2012), 166-175, doi: 10.1016/j.jlp.2011.08.004.
23. Maleki, N., Bagherifard, M. and Gholamian, M., "Application of incomplete analytic hierarchy process and choquet integral to select the best supplier and order allocation in petroleum industry", ***International Journal of Engineering***, Vol. 33, No. 11, (2020), 2299-2309, doi: 10.5829/ije.2020.33.08b.20.

Persian Abstract

چکیده

واحد تولید هیدروکسید آمونیوم (AHPU) یکی از فرآیندهای رایج و مهم موجود در بسیاری از صنایع مانند نیروگاه های نفت ، گاز ، پتروشیمی و نیروگاه ها است. با توجه به کاربرد گسترده AHPU ، ایجاد یک استراتژی نگهداری بهینه ضروری است. در این تحقیق ، برای اولین بار ، یک برنامه دقیق بازرسی سیستماتیک برای AHPU بر اساس اولویت بندی تجهیزات با استفاده از تجزیه و تحلیل خطر تجهیزات تعیین می شود. خطر تجهیزات بر اساس احتمال و نتیجه خرابی طبق استاندارد API RP 581 محاسبه می شود. بر خلاف روش های بازرسی سنتی ، که معمولاً براساس نظرات و رهنمودهای تولیدکنندگان استوار است ، چندین روش محیطی پیشنهادی روش بازرسی در نظر گرفته و تجزیه و تحلیل شد. علاوه بر این ، استراتژی های مختلف نگهداری با توجه به معیارهای مختلف مانند ایمنی ، هزینه و امکان سنجی با استفاده از فرایند سلسله مراتبی تحلیلی مقایسه شدند و در نتیجه سیاست تعمیر و نگهداری بهینه مبتنی بر ریسک برای هر گروه از تجهیزات مشخص شد. نتایج نشان داد که ، برای تجهیزات با خطر بالا و متوسط تا زیاد ، مانند مخزن ذخیره آمونیاک مایع و خطوط مربوط به آن ، روش نگهداری پیشگیرانه با دوره بازرسی ۴۸ ماه مناسب ترین روش است ، در حالی که برای خطر متوسط مانند ذخیره سازی هیدروکسید آمونیوم مخزن و تجهیزات کم خطر مانند مخزن جذب آمونیاک ، نگهداری اصلاحی به ترتیب با دوره بازرسی ۷۲ و ۹۰ ماه ، بهترین انتخاب است. سرانجام ، انجام تجزیه و تحلیل های لازم و تعیین مناسب ترین برنامه نگهداری برای هر گروه از تجهیزات با سطوح مختلف خطر در AHPU ، منجر به صرفه جویی در هزینه خاموش شدن ، هزینه بازرسی ، کاهش خرابی و افزایش ضریب اطمینان تجهیزات می شود.

IJE TRANSACTIONS C: Aspects Vol. 34, No. 9, (September 2021) 2097-2106

International Journal of Engineering

Journal Homepage: www.ije.ir

Multi-objective Optimization Design and Verification of Interior Permanent Magnet Synchronous Generator Based on Finite Element Analysis and Taguchi Method

S. R. Karimpour[a], M. R. Besmi*[a], S. M. Mirimani[b]

[a] *Faculty of Electrical and Computer Engineering, Shahed University, Tehran, Iran*
[b] *Faculty of Electrical and Computer Engineering, Babol Noshirvani University of Technology, Babol, Iran*

PAPER INFO

Paper history:
Received 15 Jan 2020
Received in revised form 11 July 2021
Accepted 23 July 2021

Keywords:
Interior Permanent Magnet Machine
Finite Element Method
Multi-objective Optimization Design
Orthogonal Matrix
Taguchi Method

ABSTRACT

This paper proposes the optimal design process of an interior permanent magnet synchronous generator (IPMSG) for wind power systems using the finite-element analysis (FEA). A multi-objective optimization design of PM generator based on Taguchi method is proposed. This paper takes into consideration as a design parameters the influence of the pole arc angle, magnet inset, magnet thickness, magnet width, stator tooth width and slot depth. The main characteristics of generator efficiency, torque ripple and output power are taken as optimization objectives. The orthogonal matrix is established according to the number of selected parameters and the level factor of each parameter. Also, FEM is used to solve the experimental matrix. As a result, an improved generator was designed and selected, which had higher maximum output power and efficiency and lower torque ripple. Finally, a prototype IPMSG was manufactured based and tested on results analysis and Taguchi method. The experimental tests were conducted to verify the validity of the proposed design process and the effectiveness of the generator and as a result, perfectly cleared the optimization design.

***doi**: 10.5829/ije.2021.34.09c.07*

NOMENCLATURE

n_n	Rated speed of the generator	Q	The number of levels
I_n	Rated current of the generator	D_{os}	The outer diameter of stator
P_n	Rated power of the generator	D_{or}	The outer diameter of rotor
η	The efficiency of the generator	g	The air-gap length between rotor and stator
NF	The Noise Factor of the Taguchi Test	l	The stack length of generator
m	The overall mean of the experimental results	A	The stator tooth width
m_η	The overall mean of the efficiency (η)	B	The stator slot depth
m_{Tr}	The overall mean of the torque ripple	C	The rotor magnet thickness
m_P	The overall mean of the output power	D	The rotor magnet width
$m_{\eta A}$	The mean of the efficiency (η) under each level of factor A	E	The rotor magnet inset
S_A	The variance of definite objective function under factor A	F	The rotor magnet spread angle
$S_{\eta A}$	The variance of the efficiency under factor A	M470-50A	The stator and rotor core material
S_{Tr}	The variance of the torque ripple under factor A	NdFeB	The permanent magnet material
S_P	The variance of the output power under factor A		

1. INTRODUCTION

Permanent magnet (PM) machines are popular and attractive for many Applications with high-efficiency, such as wind turbine systems and have more advantages than traditional machines due to their high efficiency, high power density, noticeable torque-to-current ratio, and low maintenance cost. The design of generator in wind turbine systems depends on ease of maintenance, efficiency and generator performance [1, 2]. Therefore, their design optimization process becomes more

*Corresponding Author Institutional Email: besmi@shahed.ac.ir (M. R. Besmi)

Please cite this article as: S. R. Karimpour, M. R. Besmi, S. M. Mirimani, Multi-objective Optimization Design and Verification of Interior PMSG Based on Finite Element Analysis and Taguchi Method, International Journal of Engineering, Transactions C: Aspects Vol. 34, No. 9, (2021)...2097-2106

complex as more engineering domains and constraints are involved. Design optimization of PM generators is always one of the most important topics; which has been discussed intensively in literature [3-12]. Considering the manufacturing quality of the electrical machines in practical production, robust optimization model can be developed to replace the conventional deterministic model [13, 14].

Minimization of the torque ripple is an issue for design of IPMSG. To reduce the ripple, some investigations have already been through on the rotor structure with different geometries in different poles and structures [15, 16].

Presently, many experts and researchers have presented several optimization algorithms into the multi-objective optimization design of the generator such as genetic algorithm. That may be complex, according to design structure of electrical machines [17]. To decrease the rejection rates and increase the manufacturing reliability of electrical machines, robust design optimization is always a necessity. Multi-objective optimization has become popular in this field nowadays, as design optimization, such as maximizing the torque, efficiency, power density, output power and minimizing the torque ripple, active material cost, machine weight. To achieve the multi-objective optimization of generator design and solve the different parameters, this paper pursues Taguchi method to optimal design of interior PM generator.

Taguchi method is an optimal design method that can recognize the multi-independent optimization Parameters, and it is initiated by Doctor Taguchi in the 1970s. This method is developed from quality engineering with different specific focuses on probabilistic analysis and optimization. The best alternative concerning the chosen objective metrics is selected as the solution. The Taguchi method can support the engineers to speedily discovery the optimal combination of the design variables by the minimum amount of investigates. Taguchi has the advantages of fast convergence
and high efficiency. Presently, the Taguchi method has been extensively used in the optimization of interior and surface-mounted permanent magnet machines, such as being used for improvement of efficiency and some other electromagnetic performances of the machine [18-23]. The purpose of this paper is to design and optimize IPMSG for wind power generation.

In this paper, we will deal with the efficiency and output power improvement and torque ripple reduction of a 8-pole, 27-slot, three-Phase interior permanent magnet synchronous generator (IPMSG). This aim can be effectively achieved by optimal design of the stator and rotor structure using the Taguchi and finite element methods. To get the multi-objective optimisation of generator design and solve the contrasts among different parameters, this paper adopts Taguchi method to design multi-objective optimisation of IPM generator. The Taguchi method was applied in order to evaluate and determine the geometric variables.

The optimization variables are selected such as the pole arc angle, magnet inset, magnet thickness, magnet width, stator tooth width and slot depth affect the performance of the generator. The orthogonal table of the Taguchi method was designed based on the number of geometry variables and its value levels. Multi-objective optimization was done for the main characteristics of the generator efficiency, output power and torque ripple. Finally, optimized prototype has been constructed. The finite element simulation results and the prototype experiments verify the effectiveness of the optimization method.

2. MODEL AND SPECIFICATION OF IPMSG

A three-phase Interior permanent magnet synchronous generator (IPMSG) is the object of the paper, which the geometry of the machine before improving the stator and rotor parameter is shown in Figure 1. It has buried permanent magnets. The initial values and basic parameters of the machine parameters are shown in Table 1 The stator and rotor core are composed of 0.5mm laminations of low-loss steel (M470-50A steel). The arrangement of magnets in the rotor is chosen in V shape. The stator winding adopts the fractional slot, over lapping type of double layer windings that is chosen, as it's widely used in PM machines. The machine is a small-scale prototype to be used in wind turbine application the combination of poles and slots used is 8 poles and 27 slots.

Some comparisons are made to select the best solution to meet the requirements. Efficiency, output power and torque ripple is precisely calculated by FEM and then used in Taguchi method to select the best combination as final design.

3. MULTI-OBJECTIVE OPTIMISATION DESIGN OF IPMSG

Optimisation is a systematic approach in the design of generator to make a decision about the optimum design with the appropriate goal to determine the best results such as efficiency, torque ripple, output power and all the design requirements. In this paper, Taguchi method is used to optimize the generator. This technique is an experimental design method and aims to improve product quality.

This method dramatically reduces the number of tests by using orthogonal arrays and minimizes the effects of factors that cannot be controlled, considering

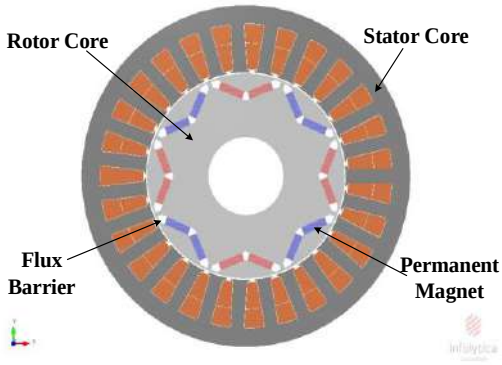


Figure 1. The geometry of the IPMSG

TABLE 1. Specifications of the IPMSG

Parameter	Symbol	Value	Unit
Rated speed	n_n	1000	r/min
Rated current	I_n	2	A
Rated Power	P_n	600	W
Pole number	p	8	––
Slot number	Q	27	--
Outer diameter of rotor	D_{or}	66	mm
Air-gap length	g	0.4	mm
Outer diameter of stator	D_{os}	109.6	mm
Stack length	l	70	mm
Tooth width	A	4.3	mm
Slot depth	B	16.5	mm
Magnet thickness	C	2.5	mm
Magnet width	D	9	mm
Magnet inset	E	5	mm
Pole arc angle	F	42	°
Stator and rotor core material	M470-50A	--	--
Permanent magnet	NdFeB	--	--

the noise factor according to the environmental conditions in the IPMSG. Furthermore, it provides a simple, efficient and systematic approach to specifying the optimum design manufacturing process.

3. 1. Optimization Variables and Objective Function The selected optimization variables are shown in Figure 2 and the optimization variables are represented as *A*, *B*, *C*, *D*, *E* and *F*, respectively. As shown in Figure 2, *A* represents the stator tooth width; *B* represents the stator slot depth; C represents the magnet thickness; D represents the magnet width; E represents the magnet inset and the angle F represents the magnet spread angle. For each parameter, five levels of influence factors are selected, and the Design Variables, Noise Factors and level values are shown in Table 2. The selected optimization objectives include the generator efficiency, torque ripple and output power. Among them, the generator efficiency and output power are expected to be better and the torque ripple is expected to be as small as possible.

3. 2. Experimental Results and Results Processing According to the six variables optimization parameters selected above and the five-level factor ranges determined by each parameter, an experimental matrix is established according to the orthogonal experiment principle of Taguchi method. The traditional single variable optimization method requires 5^6 (15625) experiments. While Taguchi method only needs 25 experiments, which can be achieved through multi-objective optimization design. The experimental $L_{25}(5^6)$ orthogonal table is established according to the number of design variables and the number of levels corresponding to each factor. The generator efficiency (η), torque ripple (Tr) and output power (P) of each experiment listed in Table 3, are calculated respectively at the 25°C and 75°C ambient temperature by applying a common finite element software-motor solve 4.1.1. The transient state solution was used to analyze and optimize the optimization objectives of each group of experiments, and the experimental matrix and solution results are shown in Table 3.

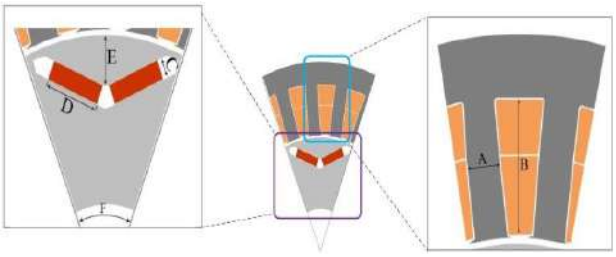


Figure 2. Design variables for Taguchi design of experiment

TABLE 2. Design Variables, Noise Factors and level values

Parameters	A (mm)	B (mm)	C (mm)	D (mm)	E (mm)	F (°)
level factor 1	4.1	15.9	2	8	4	38
level factor 2	4.2	16.1	2.1	8.5	4.5	39
level factor 3	4.3	16.3	2.2	9	5	40
level factor 4	4.4	16.5	2.35	9.5	5.5	41
level factor 5	4.5	16.7	2.5	10	6	42
Noise Factor	array pattern					
Ambient Temperature	25° C			75° C		

3. 3. Analysis of Experimental Results To analyze the influences on generator performances produced by different factors at different levels, Taguchi optimization method uses the statistical mean made by orthogonal arrays and analysis results of FEM. In the statistical analysis of the experimental results, the average is analyzed first.

3. 3. 1. Analysis of Overall Mean The overall mean of the experimental results in each column listed in 3 is analyzed, respectively, and the calculation formula is shown in Equation (1):

$$m = \frac{\sum_{i=1}^{n} m_i}{n} \quad (1)$$

where m is the overall mean of the experimental results in certain column listed in Table 3 and n is the number of experiments and m_i is the experimental result of the i experiment in certain column listed in Table 3.

The overall mean of the efficiency (η) at 25°C in Table 3 is calculated using Equation (1), and is shown as follows:

$$m_\eta = \frac{\eta_1+\eta_2+\cdots+\eta_{25}}{25} = 89.5776 \quad (2)$$

In the same approach, the overall mean of other experimental results listed in Table 3 at 2 noise factors is achieved respectively as listed in Table 4.

TABLE 3. The experimental matrix and the analysis results

Experimental number	Experimental matrix						Efficiency (%)		Torque ripple (%)		Output power (W)	
							NF1	NF2	NF1	NF2	NF1	NF2
	A	B	C	D	E	F	25° C	75° C	25° C	75° C	25° C	75° C
1	1	1	1	1	1	1	89.03	87.30	3.100	3.661	474.3	451
2	1	2	2	2	2	2	89.31	87.25	3.240	3.761	496.8	467
3	1	3	3	3	3	3	89.61	87.70	3.408	3.567	517.4	487
4	1	4	4	4	4	4	89.83	87.75	4.632	4.200	535.0	504
5	1	5	5	5	5	5	89.95	88.10	5.482	5.165	545.2	514
6	2	1	2	3	4	5	89.21	87.45	3.462	3.842	487.5	462
7	2	2	3	4	5	1	90.37	88.50	3.340	3.448	564.3	532
8	2	3	4	5	1	2	90.71	89.10	4.593	4.230	593.7	561
9	2	4	5	1	2	3	88.79	86.60	2.896	3.430	468.4	440
10	2	5	1	2	3	4	88.71	86.50	3.212	3.110	466.9	439
11	3	1	3	5	2	4	90.49	88.90	4.165	3.962	572.0	544
12	3	2	4	1	3	5	88.22	85.90	3.530	4.200	437.8	410
13	3	3	5	2	4	1	89.61	87.50	2.654	2.930	514.4	484
14	3	4	1	3	5	2	89.56	87.60	3.231	3.155	514.3	484
15	3	5	2	4	1	3	90.03	88.30	4.174	3.770	546.4	516
16	4	1	4	2	5	3	89.23	87.55	3.705	4.397	489.2	464
17	4	2	5	3	1	4	89.87	88.00	3.472	4.152	528.9	497
18	4	3	1	4	2	5	89.70	87.70	3.987	3.915	521.4	490
19	4	4	2	5	3	1	90.58	88.80	4.499	4.182	587.5	556
20	4	5	3	1	4	2	88.57	86.20	3.075	3.858	460.1	432
21	5	1	5	4	3	2	90.56	89.00	3.518	3.711	576.1	548
22	5	2	1	5	4	3	90.43	88.70	4.357	3.972	569.6	549
23	5	3	2	1	5	4	88.09	85.70	4.341	5.140	434.1	406
24	5	4	3	2	1	5	88.98	86.90	3.122	3.841	476.5	447
25	5	5	4	3	2	1	90	88.10	2.859	2.680	543.8	513

TABLE 4. Overall mean of experimental results

Operating point	m_η (%)	m_{Tr} (%)	m_P (W)
At 25° C temperature	89.5776	3.68216	516.864
At 75° C temperature	87.644	3.85116	487.88

3. 3. 2. Analysis of Mean under Each Level of Each Factor The mean of the efficiency (η) under each level of factor *A* at 25°C is set as an example to be calculated, and are shown as Equations (3)-(7).

$$m_{\eta A(1)} = \frac{\eta_1+\eta_2+\eta_3+\eta_4+\eta_5}{5} = 89.546 \quad (3)$$

$$m_{\eta A(2)} = \frac{\eta_6+\eta_7+\eta_8+\eta_9+\eta_{10}}{5} = 89.558 \quad (4)$$

$$m_{\eta A(3)} = \frac{\eta_{11}+\eta_{12}+\eta_{13}+\eta_{14}+\eta_{15}}{5} = 89.582 \quad (5)$$

$$m_{\eta A(4)} = \frac{\eta_{16}+\eta_{17}+\eta_{18}+\eta_{19}+\eta_{20}}{5} = 89.59 \quad (6)$$

$$m_{\eta A(5)} = \frac{\eta_{21}+\eta_{22}+\eta_{23}+\eta_{24}+\eta_{25}}{5} = 89.612 \quad (7)$$

In the same way, the mean of the generator efficiency, torque ripple and output power under each level of each factor at each operating point can be achieved, respectively; and the results are shown in Tables 5 and 6.

The changes of the efficiency, torque ripple and output power with the changes of the levels of each factor can be seen from Tables 5 and 6, as shown in Figure 3.

When the value taken by factor A (the stator tooth width) is larger, the efficiency of the machine is larger. The smaller the value taken by factor B (the stator slot depth), or the larger the value taken by factor *C* (the magnet thickness), the larger the increment of the efficiency and output power will be. At the same time, when the value taken by factor *D* (magnet width) is larger, the efficiency and output power will be larger. Moreover, when the value taken by factor *E* (the magnet inset), and the value taken by factor *F* (the magnet spread angle), is larger, the efficiency and output power will be less.

Further, the combinations of the level taken by each factor with the largest efficiency and output power and decrement of the torque ripple can be achieved from Tables 5 and 6, for example, at the 25°C temperature point, the combination of the level taken by each factor making the efficiency largest is *A*(5)*B*(1)*C*(5)*D*(5)*E*(1)F(1).

The combination that makes the torque ripple smallest is *A*(2)*B*(2)*C*(3)*D*(2)*E*(2)*F*(1); the combination that makes the output power largest is *A*(5)*B*(1)*C*(5)*D*(5)*E*(1)*F*(1). In the same way, the combinations of the level taken by each factor that make the efficiency and output power largest and the torque ripple smallest can be achieved at the 25°C temperature point. It can be seen from the above analysis that the combinations of the level taken by each factor that make the largest efficiency and output power, and decrement of the torque ripple are different. Consequently, it is necessary to investigate the relative importance of the effects of each design variable on the efficiency, output power and the torque ripple by the analysis of variance, and then an optimization scheme can be reached. Analysis of variance is used to evaluate the response magnitude in each parameter in the orthogonal experiment. It is used to quantify the causes of different experimental results from different parameters.

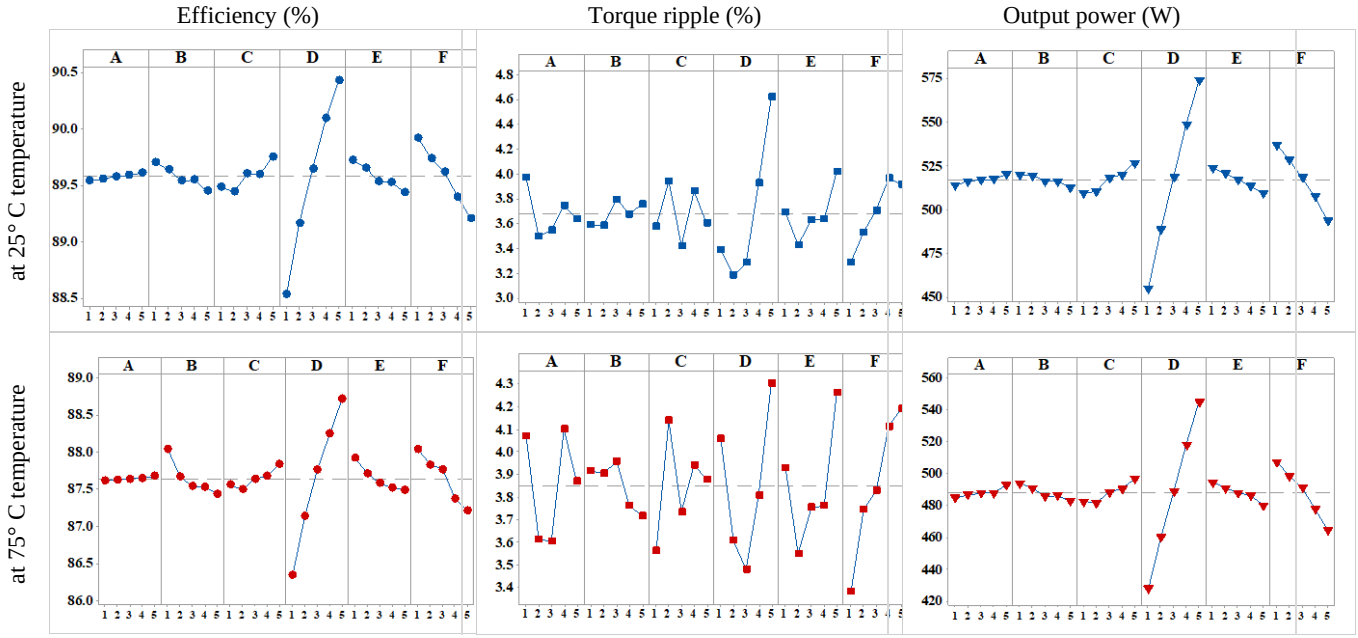
3. 4. Analysis of Variance (ANOVA) To analyze the proportions of influences on generator

TABLE 5. Mean of efficiency, torque ripple and output power under each level of each factor at 25°C temperature point

Factor	Level	m_η (%)	m_{Tr} (%)	m_P (W)
	1	89.546	3.9724	513.74
	2	89.558	3.5006	516.16
A	3	89.582	3.5508	516.98
	4	89.59	3.7476	517.42
	5	89.612	3.6394	520.02
	1	89.704	3.59	519.82
	2	89.64	3.5878	519.48
B	3	89.544	3.7966	516.2
	4	89.548	3.676	516.34
	5	89.452	3.7604	512.48
	1	89.486	3.5774	509.3
	2	89.444	3.9432	510.46
C	3	89.604	3.422	518.06
	4	89.598	3.8638	519.9
	5	89.756	3.6044	526.6
	1	88.54	3.3884	454.94
	2	89.168	3.1866	488.76
D	3	89.65	3.2864	518.38
	4	90.098	3.9302	548.64
	5	90.432	4.6192	573.6
	1	89.724	3.6922	523.96
	2	89.658	3.4294	520.48
E	3	89.536	3.6334	517.14
	4	89.53	3.636	513.32
	5	89.44	4.0198	509.42
	1	89.918	3.2904	536.86
	2	89.742	3.5314	528.2
F	3	89.618	3.708	518.2
	4	89.398	3.9644	507.38
	5	89.212	3.9166	493.68

TABLE 6. Mean of efficiency, torque ripple and output power under each level of each factor at 75°C temperature point

Factor	Level	m_η (%)	m_{Tr} (%)	m_P (W)
A	1	87.62	4.0708	484.6
	2	87.63	3.612	486.8
	3	87.64	3.6034	487.6
	4	87.65	4.1008	487.8
	5	87.68	3.8688	492.6
B	1	88.04	3.9146	493.8
	2	87.67	3.9066	491
	3	87.54	3.9564	485.6
	4	87.53	3.7616	486.2
	5	87.44	3.7166	482.8
C	1	87.56	3.5626	482.6
	2	87.5	4.139	481.4
	3	87.64	3.7352	488.4
	4	87.68	3.9414	490.4
	5	87.84	3.8776	496.6
D	1	86.34	4.0578	427.8
	2	87.14	3.6078	460.2
	3	87.77	3.4792	488.6
	4	88.25	3.8088	518
	5	88.72	4.3022	544.8
E	1	87.92	3.9308	494.4
	2	87.71	3.5496	490.8
	3	87.58	3.754	488
	4	87.52	3.7604	486.2
	5	87.49	4.261	480
F	1	88.04	3.3802	507.2
	2	87.83	3.743	498.4
	3	87.77	3.8272	491.2
	4	87.37	4.1128	478
	5	87.21	4.1926	464.6



Variable level factor
(a) Influence of parameter level factor on the efficiency

Variable level factor
(b) Influence of parameter level factor on the Torque ripple

Variable level factor
(c) Influence of parameter level factor on the output power

Figure 3. Influence of parameter level factor on target performance

performances that are produced by each factor's different levels, analysis of variance that provides a measure of confidence is conducted. The sum of squares is a measure of the deviation of the experimental data from the average value of data. ANOVA is a statistical method which is used to determine the individual interactions of all of the control factors in the experimental design. S_A generated by various factors and different levels can be calculated as Equation (8):

$$S_A = \frac{\sum_{i=1}^{Q}(m_{A(i)} - m)^2}{Q} \quad (8)$$

where S_A is the variance of definite objective function under factor *A*; $m_{A(i)}$ is the mean of certain objective function under level *i* of factor *A*; *m* is the overall mean of definite objective function; *Q* is the number of levels taken by each factor, and *Q*=5 in this paper.

Calculation for the variance of the efficiency under factor *A* at the 25° C temperature point is set as an example to explain the calculation of the variance, and the expression is shown as follows:

$$S_{\eta A} = \frac{\sum_{i=1}^{5}(m_{\eta A(i)} - m_{\eta})^2}{5} = 0.000548 \quad (9)$$

In the same approach, the variances of the efficiency, torque ripple and the output power under each factor at each operating point are calculated using Equation (8), and the results are listed in Table 7.

It can be observed from Table 7 that factor *D* has the largest impact on the efficiency, output power and the torque ripple at both of the operating points, as shown in Table 7, the efficiency is sensitive to *D, F, C, E, B* and *A*, respectively at the 25° C temperature point; *D* is the most sensitive factor; *A* has a relatively weak effect on it. Similarly, the sensitivity and proportion percent of each factor on objectives at each operating points is achievable.

3. 5. Optimization Scheme of the Improved IPMSG It can be observed from Table 5 that the combination of the level taken by each factor to make the efficiency largest is *A*(5)*B*(1)*C*(5)*D*(5)*E*(1)*F*(1), and the combination of the level taken by each factor to make the torque ripple smallest is *A(2)B(2)C(3)D(2)E(2)F(1)*, and the combination that makes the output power largest is *A(5)B(1)C(5)D(5)E(1)F(1)*, at the 25°C temperature point. in the three combinations, the level taken by factor *F* is the same and it is *F(1)*, and in the two combinations, the levels taken by factors *A, B, C, D* and *E* are the same and they are *A(5)B(1)C(5)D(5)E(1)*. It can be seen from Table 7 that the relative importance of the effects of each factor on the efficiency is DFCEBA from the largest to smallest, and the relative importance of the effects of each factor on the torque ripple is DFCEAB from the largest to smallest, and the relative importance of the effects of each factor on the output power is DFCEBA from the largest to smallest. In the same way, the combinations of the level taken by each factor to make the efficiency and output power largest and the torque ripple smallest at the 75°C temperature point are achieved respectively according to the Tables 5 and 6. Then the relative importance of the effects of each factor is achieved according to Table 7.

Finally, optimization schemes of the improved IPMSG at the 25° C and 75° C temperature point are achieved, respectively as shown in Table 8.

As a result, the best combination of factors for maximum efficiency and output power and minimum torque ripple, by the six optimization variables in the

TABLE 7. Calculation results of the variance

Factor	S_η (%)				S_{Tr} (%)				S_P(W)			
	25° C	Proportion (%)	75° C	Proportion (%)	25° C	Proportion (%)	75° C	Proportion (%)	25° C	Proportion (%)	75° C	Proportion (%)
A	0.000548	0.101	0.00042	0.049	0.028114	6.152	0.045891	14.255	4.107584	0.197	6.8576	0.341
B	0.00753	1.389	0.04458	5.087	0.007330	1.604	0.008860	2.752	7.103264	0.341	15.7216	0.783
C	0.011836	2.183	0.01350	1.541	0.037168	8.133	0.037681	11.704	40.73254	1.958	30.5056	1.519
D	0.450089	83.002	0.69906	79.759	0.285614	62.494	0.089102	27.676	1771.0805	85.114	1704.6816	84.902
E	0.010165	1.875	0.02474	2.823	0.036499	7.986	0.056585	17.576	26.295584	1.264	23.1936	1.155
F	0.06209	11.45	0.09414	10.741	0.062298	13.631	0.083822	26.037	231.51478	11.126	226.9056	11.3

TABLE 8. Optimization schemes of the improved IPMSG

	Combination (η)	Combination (Tr)	Combination (P)	Effect (η)	Effect (Tr)	Effect (P)	Final optimization scheme
Results at 25° C temperature point	A(5)B(1)C(5) D(5)E(1)F(1)	A(2)B(2)C(3) D(2)E(2)F(1)	A(5)B(1)C(5) D(5)E(1)F(1)	DFCEB A	DFCEA B	DFCEBA	A(5)B(1)C(5)D(5)E (1)F(1)
Results at 75° C temperature point	A(5)B(1)C(5) D(5)E(1)F(1)	A(3)B(5)C(1) D(3)E(2)F(1)	A(5)B(1)C(5) D(5)E(1)F(1)	DFBEC A	DFEAC B	DFCEBA	A(5)B(1)C(5)D(5)E (1)F(1)
Final optimization scheme	-	-	-	-	-	-	A(5)B(1)C(5)D(5)E (1)F(1)

final optimization scheme are *A(5)B(1)C(5)D(5)E(1)F(1)*. The specific values taken by each factor are listed in Table 9.

4. RESULT COMPARISONS AND OPTIMISED PROTOTYPE AND EXPERIMENTAL VERIFICATION

The design parameters of the IPMSG are optimized to satisfy the generator efficiency and output power and reduce the torque ripple through the proposed design process. Generator efficiency, output power and torque ripple are analyzed using FEM for the optimization model via the Taguchi method. According to the above optimization results, an Interior permanent magnet synchronous generator is designed and the prototype of the IPMSG was manufactured to validate the performances of optimized machine. Figures 4 and 5 show the components and test set of the IPMSG.

Figure 6 shows the waveform of back Emf at rated speed before and after optimization at initial and optimized model. It can be seen that after optimization, the amplitude of back EMF increased from 130.7 to 149.5 V, which increase about 14.38%.

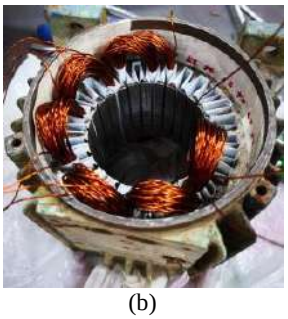
The generator is dragged to the rated speed using the prime mover, and the phase voltage is measured. The comparison between the phase voltage and the FEM are shown in Figure 7. It can be seen that the measured waveform corresponds to the analytical one.

Table 10 compares the structure parameters and output performances of the IPMSG between the initial, optimized and manufactured machine and we can see that the results improved. It can be seen that the generator efficiency and output power improves as optimization progresses, and the

TABLE 9. Values taken by each factor in the final optimization scheme of the improved IPMSG

Factor	A (mm)	B (mm)	C (mm)	D (mm)	E (mm)	F (°)
Value	4.5	15.9	2.5	10	4	38

(a)



(b)

(c)

(d)

Figure 4. Generator components. (a) Rotor and stator steel sheet, (b) stator winding, (c) components, and (d) housing

Figure 5. Test set for experiment

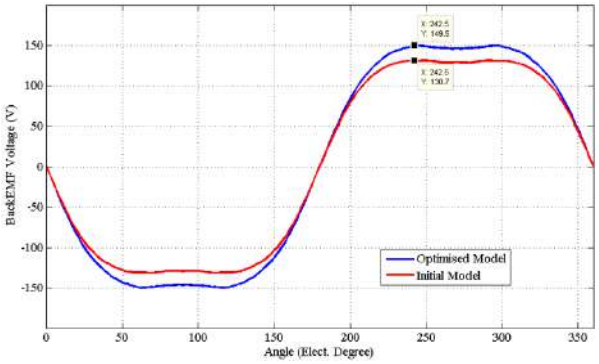
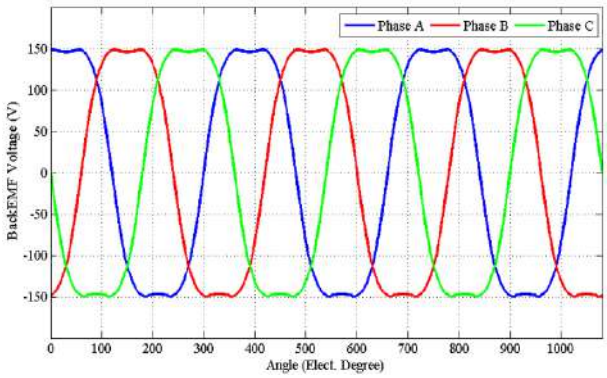
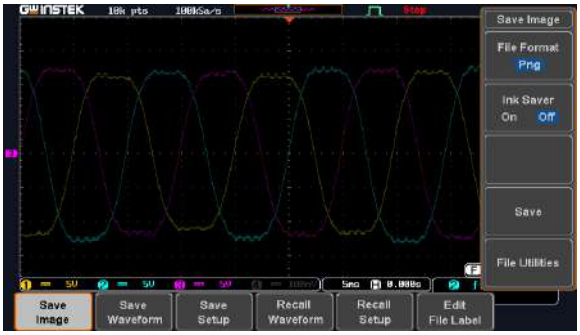


Figure 6. Comparison of results before and after optimisation at initial and optimized model

FEA and the test results are in good agreement. As well as, it can be seen that the torque ripple is reduced as the optimization design progresses. According to the proposed optimization design method, the optimized IPMSG compared to initial one improved the generator efficiency and output power about 1.797% and 15.38%, respectively, and the torque ripple is reduced by 11.72% at the rated speed and 25°C temperature point.



(a)



(b)

Figure 7. Back EMF waveform of (a) the analyzed result and (b) the experimental result

TABLE 10. Comparison of Parameters and results

Model	Parameter value						Response value		
	A (mm)	B (mm)	C (mm)	D (mm)	E (mm)	F (°)	Efficiency (%)	Torque ripple (%)	output power (W)
Initial	4.3	16.5	2.5	9	5	42	89.4	4.35	533
Optimized	4.5	15.9	2.5	10	4	38	91	3.84	615
manufactured	4.5	15.9	2	10	4	38	90.8	3.89	594

5. CONCLUSION

This paper presents a multi-objective optimization design based on the Taguchi–FEA. The main performance characteristics such as generator efficiency, torque ripple and output power are taken as the optimization objectives, and the range of parameters that have a great influence on the generator optimization objective, which is reasonably selected as optimization variables. Taguchi method is utilized to optimize the improved IPMSG at the 25° C and 75° C temperature point. The effects of each optimization variable of the improved IPMSG on the generator efficiency, torque ripple and output power is analyzed, and then the optimization schemes at each operating point are achieved.

Finally, the final optimization scheme of the improved IPMSG is achieved. The improved machine effectively increases efficiency and output power; on the other hand, it also effectively reduces the torque ripple.

Finally, a prototype 0.6 KW, 8-pole/27-slot IPMSG was manufactured based on analysis results and Taguchi method, and was tested. The experimental tests were conducted to verify the validity of the proposed design process and the effectiveness of the generator.

6. REFERENCES

1. J. H. J. Potgieter and M. J. Kamper, “Design optimization of directly grid-connected pm machines for wind energy applications”, IEEE Transactions on Industry Applications, Vol. 51, No. 4, (2015), 2949-2958. DOI: 10.1109/TIA.2015.2394506
2. Y. Duan and R. G. Harley, “Present and future trends in wind turbine generator designs”, ***IEEE Power Electronics and Machines in Wind Applications***, (2009), 1-6. DOI: 10.1109/PEMWA.2009.5208401.
3. A. Di Gerlando, G. Foglia, M. F. Iacchetti, and R. Perini, “Axial flux pm machines with concentrated armature windings: Design analysis and test validation of wind energy generators”, ***IEEE Transactions on Industrial Electronics***, Vol. 58, No. 9, (2011), 3795-3805. DOI: 10.1109/TIE.2010.2081956.
4. D. Fodorean, A. Djerdir, I.-A. Viorel, and A. Miraoui, “A double excited synchronous machine for direct drive application - design and prototype tests”, ***IEEE Transactions on Energy Conversion***, Vol. 22, No. 3, (2007), 656-665. DOI: 10.1109/TEC.2007.896279.
5. F. B. Chaaban, “Determination of the optimum rotor/stator diameter ratio of permanent magnet machines”, ***Electric Machines and Power Systems***, Vol. 22, No. 4 (1994), 521-531. DOI: 10.1080/07313569408955584.
6. L. J. Wu, Z. Q. Zhu, J. T. Chen, Z. P. Xia, and G. W. ewell, “Optimal split ratio in fractional-slot interior permanent-magnet machines with non-overlapping windings”, ***IEEE Transactions on Magnetics***, Vol. 46, No. 5, (2009), 1235-1242. DOI: 10.1109/IEMDC.2009.5075435.
7. A. Di Gerlando, G. Foglia, M. F. Iacchetti, and R. Perini, "Axial flux PM machines with concentrated armature windings: design analysis and test validation of wind energy generators", ***IEEE Transactions on Industrial Electronics***, Vol. 58, No. 9, (2011), 3795-3805. DOI: 10.1109/TIE.2010.2081956.
8. J. Tapia, J. Pyrhonen, J. Puranen, S. Nyman, and P. Lindh, "Optimal design of large permanent magnet synchronous generators", ***IEEE Transactions on Magnetics***, Vol. 49, No. 1, (2013), 642-650. DOI: 10.1109/TMAG.2012.2207907.
9. M. Moallem, A. Kiyoumarsi and M. R. Hassanzadeh, "A Novel technique on the analytical calculation of open circuit flux

density distribution in brushless permanent magnet motor", ***International Journal of Engineering, Transactions B: Applications***, Vol. 17, No. 1, (2004), 51-59. DOI: 10.5829/ije.2004.17.01b.06.

10. H. Moradi Cheshmeh Beigi, "Design, Optimization and FEM Analysis of a Surface-Mounted Permanent-magnet Brushless DC Motor", ***International Journal of Engineering, Transactions B: Applications***, Vol. 31, No. 2, (2018), 339-345. DOI: 10.5829/ije.2018.31.02b.19.
11. H. Fekri, M. A. Shamsi-Nejad and S. M. Hasheminejad, "Performance Analysis of a Novel Three-phase Axial Flux Switching Permanent Magnet Generator with Overlapping Concentrated Winding", ***International Journal of Engineering, Transactions B: Applications***, Vol. 32, No. 2, (2019), 286-295. DOI: 10.5829/ije.2019.32.02b.14.
12. A. N. Patel, B. N. Suthar, "Cogging Torque Reduction of Sandwiched Stator Axial Flux Permanent Magnet Brushless DC Motor using Magnet Notching Technique", ***International Journal of Engineering, Transactions A: Basics***, Vol. 32, No. 7, (2019), 940-946. DOI: 10.5829/ije.2019.32.07a.06.
13. Lei, G.; Zhu, J.G.; Guo, Y.G, "Multidisciplinary Design Optimization Methods for Electrical Machines and Drive Systems", Springer-Verlag: Berlin/Heidelberg, Germany, (2016), 1-6. DOI: 10.1007/978-3-662-49271-0.
14. Yao, D.; Ionel, D.M, "A review of recent developments in electrical machine design optimization methods with a permanent magnet synchronous motor benchmark study", ***IEEE Transactions on Industry Applications***, Vol. 49, No. 3 (2013), 1268-1275. DOI: 10.1109/TIA.2013.2252597.
15. N. Bianchi, S. Bolognani, D. Bon, M. D. Pre, "Rotor flux-barrier design for torque ripple reduction in synchronous reluctance and PM-assisted synchronous reluctance motors", ***IEEE Transactions on Industry Applications***, Vol. 45, (2009), 921-928. DOI: 10.1109/TIA.2009.2018960.
16. M. Sanada, K. Hiramoto, S. Morimoto, Y. Takeda, "Torque ripple improvement for synchronous reluctance motor using asymmetric flux barrier arrangement", 38th IAS Annual Meeting, Conference Record of the IEEE, (2003), 250-255. DOI: 10.1109/IAS.2003.1257510.
17. G. V. Cvetkovski, L. B. Petkovska, "Weight reduction of permanent magnet disc motor for electric vehicle using genetic algorithm optimal design procedure", ***IEEE EUROCON*** (2009). DOI: 10.1109/EURCON.2009.5167738.
18. Chang-Chou Hwang, Li-Yang Lyu, Cheng-Tsung Liu, Ping-Lun Li, "Optimal Design of an SPM Motor Using Genetic Algorithms and Taguchi Method", ***IEEE Transactions on Magnetics***, Vol. 44, No. 11, (2008), 4325-4328. DOI: 10.1109/TMAG.2008.2001526.
19. Ki-Chan Kim, Ju Lee, Hee Jun Kim, Dae-Hyun Koo, "Multiobjective Optimal Design for Interior Permanent Magnet Synchronous Motor", ***IEEE Transactions on Magnetics***, Vol. 45, No. 3, (2009), 1780-1783. DOI: 10.1109/TMAG.2009.2012820.
20. Mahdi Ashabani, Yasser Abdel-Rady I. Mohamed, Jafar Milimonfared, "Optimum Design of Tubular Permanent-Magnet Motors for Thrust Characteristics Improvement by Combined Taguchi–Neural Network Approach", ***IEEE Transactions on Magnetics,*** Vol. 46, No. 12, (2010), 4092-4100. DOI: 10.1109/TMAG.2010.2067450.
21. Chang-Chou Hwang, Chia-Ming Chang, Cheng-Tsung Liu, "A Fuzzy- Based Taguchi Method for Multiobjective Design of PM Motors", ***IEEE Transactions on Magnetics***, Vol. 49, No. 5, (2013), 2153-2156. DOI: 10.1109/TMAG.2013.2242854.
22. Kim, S.-I.; Lee, J.-Y.; Kim, Y.-K.; Hong, J.-P.; Hur, Y.; Jung, Y. H., "Optimization for reduction of torque ripple in interior permanent magnet motor by using the Taguchi method", ***IEEE Transactions on Magnetics***, Vol. 41, (2005), 1-5. DOI: 10.1109/TMAG.2005.846478.
23. J. Zhu, S. Li, D. Song, Q. Han, J. Wang, G. Li, "Multi-objective optimization design of air-cored axial flux PM generator", ***IET Electric Power Applications,*** (2018), 1390-1395. DOI: 10.1049/iet-epa.2018.5092.
24. S. R. Karimpour, M. R. Besmi, S. M. Mirimani, "Optimal design and verification of interior permanent magnet synchronous generator based on FEA and Taguchi method", ***International Transactions on Electrical Energy Systems,*** (2020), e12597, 1-15. DOI: 10.1002/2050-7038.12597.

Persian Abstract

چکیده

در این مقاله فرآیند طراحی بهینه یک ژنراتور سنکرون مغناطیس دائم با آهنربای داخلی (IPMSG) برای کاربرد در سیستم‌های بادی به کمک روش تجزیه و تحلیل اجزای محدود (FEA) ارائه شده است. روش طراحی بهینه‌سازی چند هدفه ژنراتور مغناطیس دائم بر اساس روش تاگوچی پیشنهاد گردیده است. این مقاله تأثیر زاویه گام قطب آهنربا، فاصله قرارگیری آهنربا تا فاصله هوایی، ضخامت آهنربا، عرض آهنربا، عرض دندانه استاتور و عمق شیار را به عنوان پارامترهای طراحی در نظر گرفته است. مشخصه‌های اصلی بازدهی ژنراتور، انحراف گشتاور و توان خروجی به عنوان اهداف بهینه‌سازی در نظر گرفته شده است. ماتریس متعامد تاگوچی با توجه به تعداد پارامترهای انتخاب شده و ضریب سطح هر پارامتر و روش اجزای محدود، به منظور حل ماتریس آزمایش‌های تاگوچی به کار گرفته شده است. در نتیجه، یک ژنراتور بهینه شده، طراحی و انتخاب گردید، که دارای ویژگی‌های حداکثر توان و بازدهی و انحراف گشتاور پایین‌تر می‌باشد. در پایان، یک ژنراتور نمونه آزمایشگاهی بر اساس نتایج آنالیز و روش تاگوچی ساخته شد و مورد آزمایش قرار گرفت. تست‌های آزمایشگاهی به منظور تأیید صحت فرآیند طراحی پیشنهادی و کارایی ژنراتور مزبور انجام شد و در نتیجه تائیدی بر طراحی بهینه ژنراتور مزبور می‌باشد.

International Journal of Engineering

Journal Homepage: www.ije.ir

Soft System Methodology: Project vs Local Community Interests in Project Social Conflict

H. B. Sanggoro*, S. W. Alisjahbana, D. Mohamad

Civil Engineering Department, University of Tarumanagara Jakarta, Indonesia

PAPER INFO

Paper history:
Received 2 July 2021
Received in revised form 21 July 2021
Accepted 24 July 2021

Keywords:
Social Conflict
Middle Way Interests
Soft System Methodology
Local Community
Environment And Social Framework

ABSTRACT

Social conflict in projects is an uncompromising representation of differences in interests. Each project receives different claims from stakeholders based on their conditions. Project interest depends on the target level and goals to be achieved. Subsequently, the local community surrounding the project transforms their interests based on economic, social, cultural, and environmental expectations. This study aims to reveal factors of interests between project and local community on the occurrence of social conflict in projects. Soft System Methodology and Customers, Actors, Transformation process, Weltanschauung/world view, Owner, and Environmental constraints (CATWOE) were used to identify root cause and factors of interests in the social conflict. Through library research and previous relevant researches, it was revealed that there are eight factors of interests which cause social conflict in projects. It is recommended that future researches should use factors in formal system methodology (FSM) to model the prediction of social conflict in projects based on different interests between project and local community. The results are useful to compose the standard environment and social framework (ESF) as a "middle way interests" concept to reduce social conflict in projects. By using the correct ESF, interests and conflicts can be managed well so that the benefits can be enjoyed by the community and project itself. Finally, the results of this study can be used to formulate indices of social conflicts in construction projects.

***doi**: 10.5829/ije.2021.34.09c.08*

1. INTRODUCTION

Indonesia as one of the G20 members and one of the largest GDP countries group in the world has its own challenges to be able to maintain and even increase its GDP amid the increasingly tough world economic competition. In order to achieve great level of economic growth, based on the Global Infrastructure Hub (GI Hub) report, Indonesia still needs to develop its infrastructure sector to support the economic growth. According to Global Infrastructure Outlook, the projection of the needs of infrastructure development in Indonesia reaches 1.7 trillion US dollar by 2040, or nearly 1.9 times of Indonesia's current GDP. Meanwhile, according to the 2019 Global Competitiveness Report, the quality of road infrastructure in Indonesia obtains an index value of 4.2 out of 52.6 and is in the 50th position out of 141 countries. In 2021 State Budget (APBN), the Ministry of Public Works and Housing (PUPR) is entrusted with managing 150 trillion rupiah, 25% increase compared to 2020. However, World Bank through Infrastructure Sector Assessment Program, still places the quality of infrastructure in Indonesia on average below the BRICS (Brazil, Russia, India, China, and South Africa) and ASEAN. This certainly should get serious attention from construction actors in Indonesia in order to be able to catch up with other countries. Such crucial position for quality and quantity of infrastructure for economic growth should be of greater concern for all stakeholders in order to create various breakthroughs in increasing resource capacity and quality of work in the construction sector. Apart from that, social impacts caused at construction stage also needs to be considered significant. Nowadays, Indonesia is facing the fact that

*Corresponding Author Institutional Email: heru.328201007@stu.untar.ac.id (H. B. Sanggoro)

Please cite this article as: H. B. Sanggoro, S. W. Alisjahbana, D. Mohamad, Soft System Methodology: Project vs Local Community Interests in Project Social Conflict, International Journal of Engineering, Transactions C: Aspects Vol. 34, No. 9, (2021) 2107-2115

conflicts that occur in the infrastructure sector is rising. In 2018, there were 300 conflict cases in 16 provinces as reported by the Foundation of the Indonesian Legal Aid Institute (YLBHI). Moreover, 94 out of 208 agrarian conflicts were in the infrastructure sector.

Social conflict that occurs certainly disturbs the project performance in general. Setianto [1] suggested that in addition land conflict, conflict in projects may occur due to the interests in the economic and social aspect. The difference of interests and expectations eventually trigger the conflict in the implementation of the project. Chan and Oppong [2] in their study stated that each stakeholder has their own expectations, some of whom are supportive and some other potentialy disrupt the project. The different interests of stakeholders must be managed properly. According to Project Management Body of Knowledge (PMBOK) [3], the first step is to identify the factors of interests in the project. The factors of project interest must be analyzed and revealed to determine appropriate compromising steps in order to minimize potential conflicts in the project.

Soft System Methodology is a method used to reveal and analyzed influencing factors of interests in social conflict in projects. Soft system methodology approach was used to develop the structured way of thinking from social conflicts. Referring to Eden and Ackermann [4], several methods have been collaborated in developing and revealing problem situations, such as Soft System Methodology (SSM), Strategic Choice Approach (SCA) and Strategic Options Development and Analysis (SODA). As Checkland [5] explained, Soft System Thinking examines how problematic situations should be understood and discussed so that the problems encountered in the research can be revealed more clearly, easily, and provide useful insights. This article explains the use of Soft System Methodology as a system thinking in revealing factors of interests that affect social conflict in projects. Rich pictures weret used in the analysis of problems through the approach of root cause analysis and CATWOE.

In recent years, the construction sector in Indonesia has faced social conflict due to the increasing development. Meanwhile, regulations that govern social and environmental protection remains sectoral. Therefore, those regulations are difficult to be enforced and understood by project actors in order to develop the best strategy to mediate different interests using an appropriate framework.

This study aims to find factors affecting the conflict of interests between project actors and the sorrounding community using Soft System Methodology. Furthermore, this study also aims to reveal the factors of interests needed to formulate a standard environmental and social protection framework in further research. The current model can be used as a standard framework to estimate the possibility of social conflict in projects in Indonesia based on the existing factors of interests sorrounding the project. Therefore, project actors can accurately predict the possibility of conflict and set a minimum framework needed to minimalize the impact of the conflict to the project.

2. LITERATURE REVIEW

Chan and Oppong [2] stated that conflict between the affected community and the project will occur if the project's expectations cannot be realized, in terms of both quantity and quality. This statement is confirmed by Omenge et al. [6] who stated that accumulated interests sometimes clash among the members of the project team itself. Therefore, in addition of conflict with external team, conflict can occur within the team as a result of interaction among interdependent people who perceive incompatible goals and interference from each other in achieving those goals. In general, Wu et al. [7] concluded that conflict will affect the project performance.

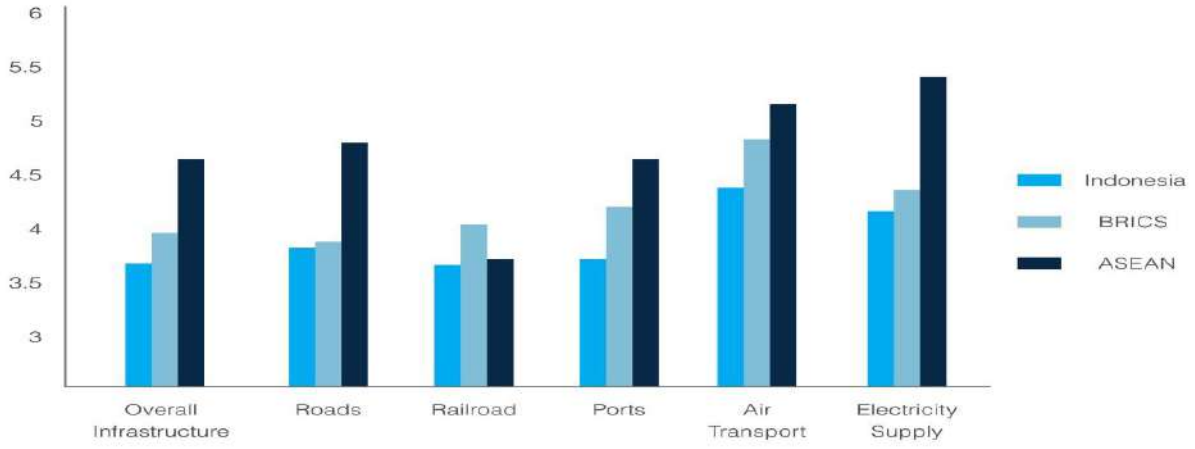


Figure 1. Indonesian infrastructure quality index

Moreover, Al-Sibaie et al. [8] found that social and internal conflicts mostly influenced the project performance. To reduce potential conflicts in the project, Vaux and Kirk [9] suggested to investigate the factors of conflict and discover the methods to mitigate the effect of the conflict.

Project stakeholders are categorized into two, namely internal and external, which are attributed according to their own interests to the project. The interests of both stakeholders interests are in relation to their position and perception on political, economic, and social system [3]. Wei et al. [10] mentioned that there are three aspects of sustainable development, namely (a) social, (b) environmental, and (c) economic. However, Chan [2], Luckmann and Färber [11] and Xiahou et al. [12] agreed that these three aspects are also factors that influence the interests or occurrence of conflicts, that is, stakeholder perceptions toward the project that affect the project performance, including the potential to create project social conflicts. The different conceptions of social, environmental, and economic conditions are factors that often occur in each country in the project implementation. However, apart from those factors, cultural differences are also important and influential factors in the project [11, 13, 14]

Internal stakeholders and business entity in the project also have their own interests. Target performance success as an objective of the project has a strong relationship with the conflicts that occur [7]. Likewise, Vaux and Kirk [9] and Min et al [15] found that project characteristics affect the conflict of the project. The higher complexity of the project, the higher the possibility of conflicts occuring in the project. Similarly, project as a business entity has the objective of corporate business. The project team must fulfill the company's organizational values, which include marketing, management, and business aspects [16, 17]. Another interest that exists in internal project is the interest of individual and team who have different objectives, goals, and motivations, which can trigger debates and conflicts in project implementation [8, 16, 18]. This is corroborated by Taghizadehalvandi and Ozturk [19] who suggested that personal satisfaction affects the project satisfaction as a whole. The interests of each stakeholder need to be managed properly through compromise and negotiation. Jang et al. [20] argued that for every difference of interest between the parties, there is always intersection that can be used as a negotiation opportunity, which is known as Zone of Possible Agreement (ZOPA) concept. According to World Bank [21], the area of negotiation opportunity in this study is referred to as compromise of interests in the concept of environmental and social framework.

Several pevious researches on conflicts have been conducted, but are limited to the influence of factors of interests on the conflict. Furthermore, particularly in Indonesia, researches on conflicts have not been widely conducted. Previous studies, as mentioned beforehand and demonstrated in the following Table 1, have provided significant information regarding factors and relationships of conflicts and projects. However, the integration between the influence of factors of interests and conflicts needs to be improved by measuring the capability of the existing framework of environmental and social protection. By acknowledging that capability, project actors can easily identify, mitigate, and plan the appropriate strategy to minimize conflicts. In the case of conflicts in Indonesia as the object of the current study, indices and statistical data were employed to measure the level of community interest within a region.The following is a table elaborating previous researches on conflict in projects.

3. METHODOLOGY

This study employed library research and comparative analysis from previous researches. Several previous researches found that influencing factors caused social conflict in projects during the construction implementation. Subsequently, problem analysis was carried out using CATWOE analysis and Soft System Methodology model approach to reveal the cause of social conflict in projects.

TABLE 1. Empirical study of project conflict

Authors	Subject Discussion
Hartono et al. [22], Park et al. [23], Riley and Ellegood [24], Wang and Xiang [25], Wu et al. [7],	Examined the relationship between factors of conflict and project performance
Caputo [26], Jang et al. [20], Omenge et al. [6]	The concept of conflict resolution, elaborating the influencing factors using several concepts such as ESIA and
Xiahou et al. [12]	Evaluation of factors influencing social performance in construction projects
Min et al. [15]	Established a framework for conflict analysis by modetaring project characteristics and intervention from the government authority
Xue and Xiang [27]	The relationship of risk factors and risky events (mass action) which affect social instability in the project
Celik et al. [28], Chen et al. [29]	Minimalizing social impact and estimating social cost as a result of project activity

Soft System Methodology is a systemic (not systematic) methodology; the focus is on the whole, not the parts. As a system-based methodology for dealing with real world problems, Mehregana et al. [30] stated that Soft System Methodology enables analysts and participants to understand different perspectives on the situations and problems solved through learning rather than replacing current situations with escalating espoused ideals. Soft System Methodology method can be very helpful and facilitate the overall research of a system of social conflict in projects. The problem approach using Soft System Methodology is carried out, in general, using a search process consisting of seven stages as illustrated in Figure 2. Thus, it is necessary to identify the crucial characteristics of the decision situation, define the scope and boundaries of the analysis, identify the stakeholders involved as well as their main motivations and goals, and understand what actions can be taken [31].

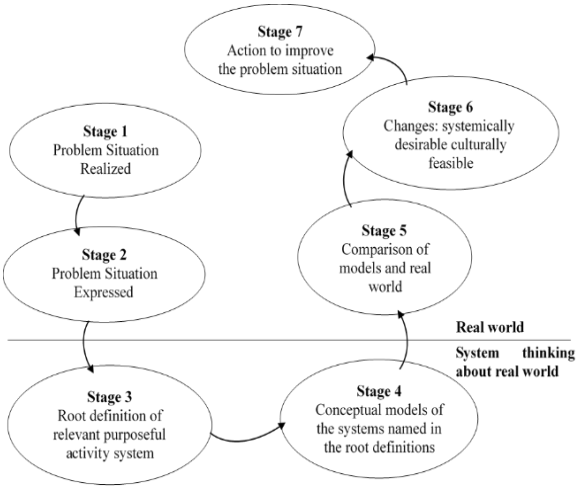


Figure 2. Soft system methodology stages

TABLE 2. CATWOE definition

	Definition
C	*Client* – the immediate beneficiaries or victims of the system results.
A	*Actors* – the participants in the transformation, i.e. those who carry out activities within the system.
T	*Transformation* – the core of the human activity system, in which some inputs are converted into outputs and given to the clients. Actors play a role in this transformation process.
W	*Weltanschauung (world view)* – the perspective or point of view that makes sense of the root definition being developed
O	*Owner* – the individual or group responsible for the proposed system. He/she has the power to modify or even stop the system, overlapping other system actors.
E	*Environmental constraints* – the human activity systems work under some constraints imposed by the external environment, as legal, physical, or ethical constraints.

The most common strategy to describe problem is using rich picture. Customers, Actors, Transformation process, Weltanschauung/world view, Owner, and Environmental constraints (CATWOE) was employed to analyze and identify problem areas, achievement, and goals of the proposed models.

4. RESULTS

A project starts when the contract is signed by the owner and contractor as the project implementer. The contract results in an agreement on achievement targets that must be mutually agreed upon [3]. The contractor will subsequently appoint a team assigned to complete the project and carry out the company's vision and missions as a business entity. This will have impact on the enactment of SOPs and targets which in parallel will create interests in the project team itself. Meanwhile, there are local communities that are affected by project activities. These conditions will underlie the existence of community expectations for activities around them, including construction activities. These expectations are then transformed into public/community interests towards the project [2]. These interests, if not managed properly, will create social conflicts in project. These interests must be solved, which in the context of this research is referred to as the compromise of interests. With good compromise, it is expected that conflicts can be avoided so that it will benefit the community and the project [20, 32, 33]. Figure 3 is a rich picture that describes the interests and conflicts in project.

Problems of social conflict in a project can affect profit losses and occur for a long time period. As in Jatigede Dam Construction Project, West Java Province, Indonesia, Setianto [1] explained that escalation and de-escalation of conflict occurred since 1982, which accounts for three decades. Social conflict in projects will result in the increasing social cost that should be spent by the project [29] and can influence the social instability in the surrounding local community [28]. To avoid of social conflict in projects, it is important to identify of the area of problems, mitigate potential conflicts, and determine goals and expected changes. Based on the aforementioned factors, identification of root cause of social conflict in this research was employed using Soft System Methodology and CATWOE method as follows.

The above CATWOE analysis (Table 3) explains the root cause conditions faced in managing social conflict in projects. The structure of CATWOE assisted in developing the conceptual research model which is the basis for determining objectives of Formal System Model. Table 4 below describes the factors of interests that influence conflict in projects.

TABLE 3. CATWOE analysis in project social conflict

	Definition
C –Customers/ Clients	Owner (Ministry of Public Work and Housing)
A –Actors	Project Team and Local Community
T –Transformation	How to compromize interests between Project Team and Local Community and what standards of environment and social framework needed to reduce potential social conflict in infrastructure project
W - Welltanschauung/ World Wide	Project Team will achieve their goals in profit and performance properly. Local affected community reaches their benefit in terms of social, economic, environment, and culture aspects. Owner gets the project meet their expectation.
O –Owner	Ministry of Public Work and Housing
E –Environmental Constraints	Various Indonesian Decree Laws and regulations in environment and social safeguards and there is no integrated standard environment and social framework.

Root cause, problem areas, and mitigation of potential conflicts that have been described beforehand must be developed within the Soft System Methodology framework. Composing problems were intended to determine objectives and needs for appropriate social conflict management.

The Soft System Methodology model indicates that the final objective of this research is to reduce potential conflict by using predictive model of interest level in

TABLE 4. Interests factors related to project conflicts

Factors	Authors [Ref.]
Project performance	Al-Sibaie et al. [8]; Meng et al. [13]; Molwus et al. [34]; Nguyen and Watanabe [35]; Taheri et al. [36]; Turner and Lecoeuvre [17]; Vaux and Kirk [9]; Wang and Xiang [25]; Wei et al. [10]; Wu et al. [7]
Project characteristic	Min et al. [15]; Molwus et al. [34]; Omenge et al. [6]; Vaux and Kirk [9]; Wu et al. [7];
Corporate management and business	Nguyen and Watanabe [35]; Panahi et al. [16]; Turner and Lecoeuvre [17]; Wang and Xiang [25]
Personal and team interests	Al-Sibaie et al. [8]; Nguyen and Watanabe [35]; Panahi et al. [16]; Taghizadehalvandi and Ozturk [19]
Social condition	Al-Sibaie et al. [8]; Çelik [28]; Chan [2]; Magsi et al. [33]; Molwus et al. [34]; Nguyen and Watanabe [35]; Puck et al. [18]; Wang and Xiang [25]; Wei et al. [10]; Xiahou et al. [12]; Xue et al. [27]
Economic condition	Celik et al. [28], Chen et al. [29]; Magsi et al. [33]; Wei et al.[10]; Xiahou et al. [12];
Environmental condition	Çelik [28]; Chan [2]; Magsi et al. [33]; Min et al. [15]; Wei et al. [10]; Xiahou et al. [12]
Culture condition	Lückmann [11]; Meng et al. [13]

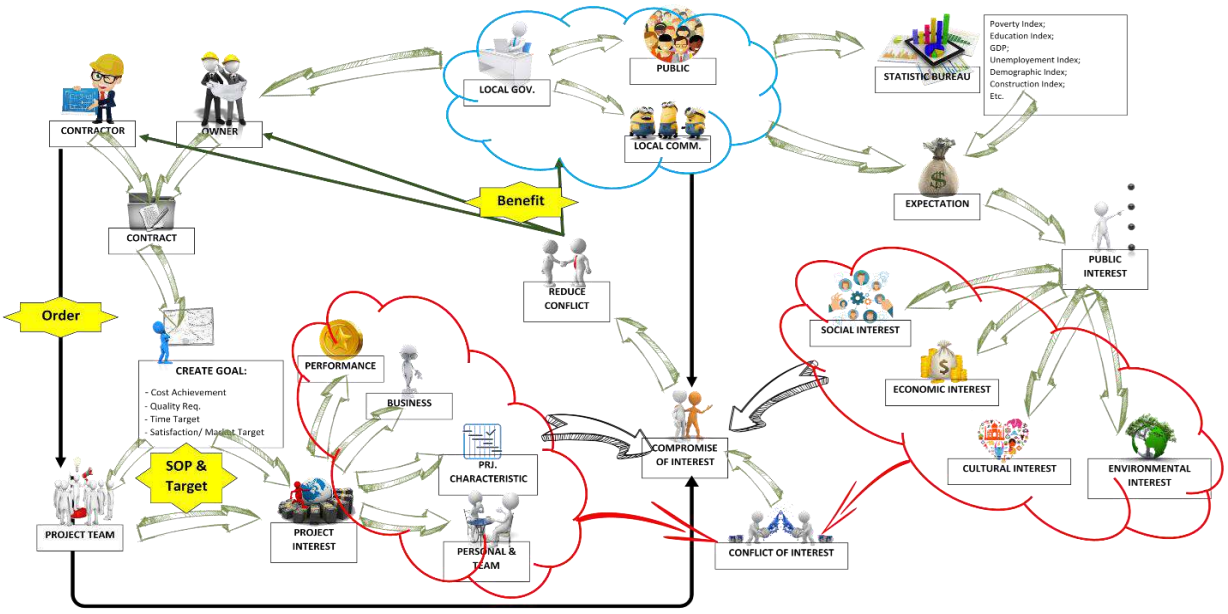


Figure 3. Rich picture : How interest and conflicts arise in project

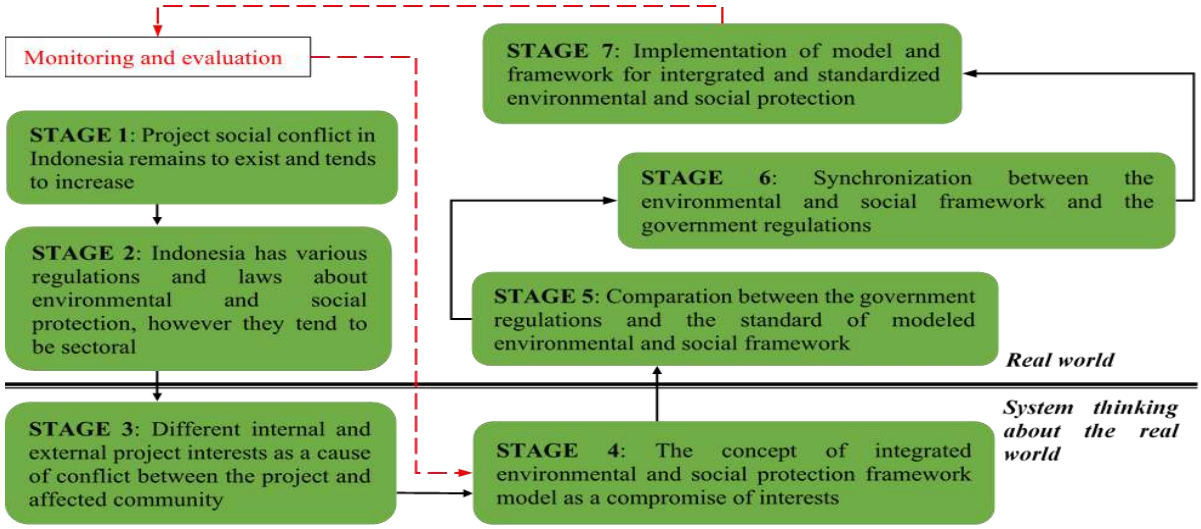


Figure 4. Soft System Methodology model of social conflict in projects

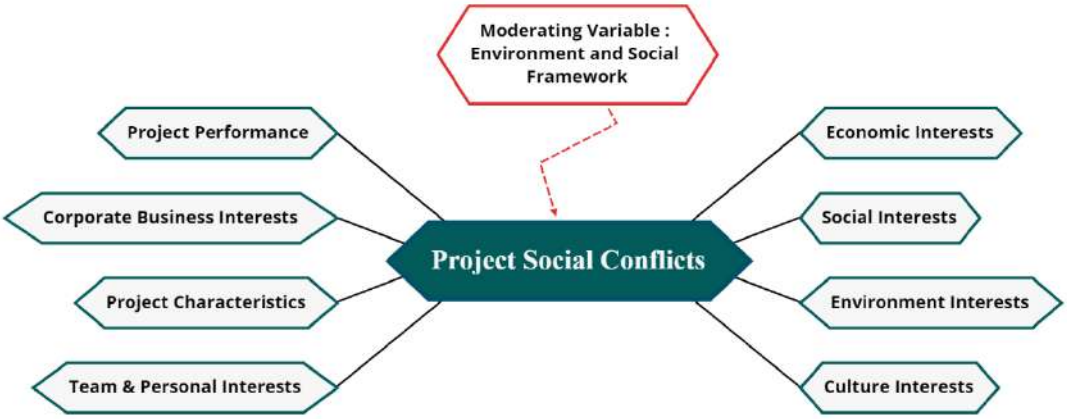


Figure 5. Factors of interests related to social conflict in projects

each project area. In addition, the model can also estimate the need for appropriate environmental and social framework standards according to interest conditions. Composing this standard framework is needed to facilitate the project in measuring and testing interests conditions around the project site.

Figure 4 above describes the stages in Soft System Methodology based on conditions and problems that have been identified and analyzed as follows:

Stage 1. Social conflict in projects still occurs with an increasing trend based on the data from YLBHI. Witrianto [31] states that in the last three decades, there has been a rapid change in public perceptions of the impacts they receive as a result of projects. These changes encouraged the community to protest and fight, which resulted in social conflict in the project.

Stage 2. There is no standard framework and model that can be used to predict potential conflicts due to the influence of interests around the project. Regulations on environmental and social protection are still sectoral.

Stage 3. The differences of interests between the internal of the project and the affected local communities as the cause of the conflict need to be identified and analyzed appropriately. Based on previous researches, there are at least eight main factors of interest of each party that have the potential to conflict, namely project performance targets, project characteristics, company business interests, personal interests, and project teams. Meanwhile, the community has interests in economic, social, environmental, and cultural aspects.

Stage 4. To minimize the potential for social conflicts in the project due to differences of interests between the internal of the project and the affected communities, it is necessary to develop an integrated environmental and social framework. This integrated standard framework is a concept of compromise of interests to reduce the impact of social risks that cause conflict in the project.

Stage 5. An integrated environmental and social framework model as a concept of compromise of interests is the middle way of Indonesian regulations regarding environmental and social protection which are still sectoral.

Stage 6. The adoption of the World Bank's environmental and social framework using a parameter measurement approach from the prevailing regulations in Indonesia to create an appropriate standard framework.

Stage 7. Implementation of the model on the project to predict potential social conflicts using the factors of interest that affect the project. With accurate predictions, the project can develop the best strategy in managing project interests to minimize potential conflicts.

Monitoring and evaluation: i.e., changes in conditions of interest in the project need to be evaluated and monitored regularly. Changes in social, economic, and political conditions, for example, can encourage changes and shifts in expectations and interests in the local community that affected the project Likewise, changes in economic, political, and monetary conditions have the opportunity to change the expectations of the contractor as a business entity which depends on the stability of national and local economic and political conditions. These changes will require adjustments to the standard framework that has been established. The results of this adjustment will be the basis for the implementation of the model to be implemented afterwards.

Figure 5 is the construction of model based on the factors of interests that influence social conflict in projects. Based on the identification and analysis conducted, the internal of the project's interests that affect the conflict consisted of factors of interest to project performance, company business interests, project characteristics, and the interests of the team and its members. Meanwhile, the interests of local communities were determined by their interests in economic, social, environmental, and cultural aspects. These eight factors are factors that influence the occurrence of conflict in the project. The greater the level of interests, the higher the possibility of conflict. The environmental and social framework was used as a moderator which was expected to reduce the power of interests in creating conflict. The measurement of public/local community interest employed indices and data published by the National Statistics Agency and Ministry/State Agency authorized in environmental and social protection. However, to measure the level of internal interests of the project, the level of the target assigned to the project was used.

5. CONCLUSION

One of the main factors that causes conflict is the difference of interests between the project and the affected local communities. By developing a conflict prediction model through an environmental and social framework approach, the author hopes to be able to facilitate project actors in predicting potential social conflicts in projects and preparing appropriate prevention steps and strategies. The results of this study can be concluded as follows:

- The model built made it easier for project actors to manage the differences of interests between the interests of the internal of the project and the interests of the local community affected by the project.
- Efforts to minimize conflict and its impacts can be more easily planned by using an appropriate standard framework as the concept of "middle way interests".
- Composing standardized and integrated framework as "middle way interests" assisted the project in identifying potential conflicts and predicting appropriate framework according to the conditions of interests in the project area.
- Eight factors were identified and analyzed in the context of social conflict management of the project. Those factors can be used in further research with Formal System Model using the data obtained from infrastructure projects in Indonesia. Environment and social framework will be used as moderating variables in the developed Formal System Model.
- Measurement of economic, social, environmental, and cultural conditions employed the data and statistical indices from the Indonesian Central Bureau of Statistics (BPS) and authorized government agencies. The data and statistical indices are research novelties which makes it easier for the model to be implemented to measure the level of public/local community interests. Project performance, corporate business interests, project characteristics, and personal and team interests employed the data from each project as research samples.

The results of this study are expected to contribute to better conflict management in the project. In addition to the use of the Formal System Model, the results of this study can be used to formulate a social conflict index formula in construction projects in future research.

6. REFERENCES

1. Setianto, S. "Social Conflict In Infrastructure Development Case Study Jati Gede Dam." ***Jurnal Sosek Pekerjaan Umum,*** Vol. 6, No. 3, (2014), 140-221. http://jurnalsosekpu.pu.go.id/index.php/sosekpu/article/view/10
2. Chan, A. P.C. and Oppong, G. D. "Managing the expectations of external stakeholders in construction projects." ***Engineering, Construction and Architectural Management,*** Vol. 24, No. 5, (2017), 736-756. https://doi.org/10.1108/ECAM-07-2016-0159

3. Project Management Institute, "A Guide to the Project Management Body of Knowledge (PMBOK)." Sixth Edition, Project Management Institute, Inc., Newtown Square-Pennsylvania, 2017.
4. Eden, C. and Ackermann, F. "Where next for problem structuring methods?." ***Journal of the Operational Research Society,*** Vol. 57, No. 7, (2006), 766-768. https://doi.org/10.1057/palgrave.jors.2602090
5. Checkland, P. "Soft systems methodology: a thirty year retrospective." ***Systems Research and Behavioral Science,*** Vol. 17, (2000), S11-S58. https://doi.org/10.1002/1099-1743(200011)17:1+<::AID-SRES374>3.0.CO;2-O
6. Omenge, P. M., Obwoyere, G. O., Eshiamwata, G. W., Makindi, S. M. and Nathwani, J. "Environmental and Social Impact Assessment Procedural Steps That Underpin Conflict Identification: Reference to Renewable Energy Resource Development In Kenya." ***International Journal of Energy Production and Management,*** Vol. 5, No. 2, (2020), 157-174. http://dx.doi.org/10.2495/EQ-V5-N2-157-174
7. Wu, G., Zhao, X. and Zuo, J. "Effects of inter-organizational conflicts on construction project added value in China." ***International Journal of Conflict Management,*** Vol. 28, No. 5, (2017), 695-723. https://doi.org/10.1108/IJCMA-03-2017-0025
8. Al-Sibaie, E. Z., Alashwal, A. M., Abdul-Rahman, H. and Zolkafli, U. K. "Determining the relationship between conflict factors and performance of international construction projects." ***Engineering, Construction and Architectural Management,*** Vol. 21, No. 4, (2014), 369-382. https://doi.org/10.1108/ECAM-03-2014-0034
9. Vaux, J. S. and Kirk, W. M. "Relationship Conflict in Construction Management: Performance and Productivity Problem." ***Journal of Construction Engineering and Management***, Vol. 144, No. 6, (2018), 04018032. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001478
10. Wei, H. H., Liu, M., Skibniewski, M. J. and Balali, V. "Conflict and consensus in stakeholder attitudes toward sustainable transport projects in China: An empirical investigation." ***Habitat International,*** Vol. 53, (2016), 473-484. https://doi.org/10.1016/j.habitatint.2015.12.021
11. Lückmann, P. and Färber, K. "The impact of cultural differences on project stakeholder engagement: a review of case study research in international project management." Procedia Computer Science, Vol. 100, (2016), 85–94. https://doi.org/10.1016/j.procs.2016.09.127
12. Xiahou, X., Tang, Y., Yuan, J., Chang, T., Liu, P. and Li, Q. "Evaluating Social Performance of Construction Projects: An Empirical Study." ***Sustainability,*** Vol. 10, No. 7, (2018), 2329. https://doi.org/10.3390/su10072329
13. Meng, J., Yan, J. and Xue, B. "Exploring Relationships between National Culture and Infrastructure Sustainability Using QCA." ***Journal of Construction Engineering and Management***, Vol. 144, No. 9, (2018), 04018082. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001463
14. Sanggoro, H. B., Widyaningsih, N. and Bintoro, B. P. K. "Analysis influence factors of domination, competency and interpersonal skill in the stakeholder interaction to infrastructure project success." International Journal of Engineering & Technology, Vol. 9, No. 1, (2020), 164-174. http://dx.doi.org/10.14419/ijet.v9i1.30153
15. Min, J. H., Jang, W., Han, S. H., Kim, D. and Kwak, Y. H. "How Conflict Occurs and What Causes Conflict: Conflict Analysis Framework for Public Infrastructure Projects." ***Journal of Management in Engineering***, Vol. 34, No. 4, (2018), 04018019. https://doi.org/10.1061/(ASCE)ME.1943-5479.0000625
16. Panahi, B., Moezzi, E., Preece, C. N. and Zakaria, W. N. W. "Value conflicts and organizational commitment of internal construction stakeholders." Engineering, Construction and Architectural Management, Vol. 24, No. 4, (2017), 554-574. https://doi.org/10.1108/ECAM-01-2016-0006
17. Turner, J. R. and Lecoeuvre, L. "Marketing by, for and of the project: project marketing by three types of organizations." ***International Journal of Managing Projects in Business,*** Vol. 10, No. 4, (2017), 841-855. https://doi.org/10.1108/IJMPB-10-2016-0080
18. Puck, J. F., Neyer, A-K. and Dennerlein, T. "Diversity and conflict in teams: a contingency perspective." ***European Journal of International Management,*** Vol. 4, No. 4, (2011), 417-439. http://dx.doi.org/10.1504/EJIM.2010.033610
19. Taghizadehalvandi, M. and Ozturk, Z. K. "Multi-objective Solution Approaches for Employee Shift Scheduling Problems in Service Sectors." ***International Journal of Engineering, IJE Transactions C: Aspects,*** Vol. 32, No. 9, (2019), 1312-1319. https://dx.doi.org/10.5829/ije.2019.32.09c.12
20. Jang, W., Yu, G., Jung, W., Kim, D. and Han, S. H. "Financial Conflict Resolution for Public-Private Partnership Projects Using a Three-Phase Game Framework." ***Journal of Construction Engineering and Management,*** Vol. 144, No. 3, (2017), 05017022. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001442
21. World Bank, "The WORLD BANK : Environmental and Social Framework," ***International Bank for Reconstruction and Development/The World Bank,*** Washington DC, 2017.
22. Hartono, B., Dzulfikar, L. and Damayanti, R. "Impact of Team Diversity and Conflict on Project Performance in Indonesian Start-Ups," ***Journal of Industrial Engineering and Management,*** Vol. 13, No. 1, (2020), 155-178. http://dx.doi.org/10.3926/jiem.3037
23. Park, C. Y., Han, S. H., Lee, K. and Lee, Y. M. "Analyzing Drivers of Conflict in Energy Infrastructure Projects: Empirical Case Study of Natural Gas Pipeline Sectors," ***Sustainability***, Vol. 9, No. 11, (2017), 2031. https://doi.org/10.3390/su9112031
24. Riley, J. M. and Ellegood, W. A. "Relationship conflict, task conflict and teams' transactive memory systems," ***International Journal of Educational Management,*** Vol. 34, No. 3, (2020), 626-640. https://doi.org/10.1108/IJEM-01-2019-0003
25. Wang, Y. and Xiang, P. "Investigate the Conduction Path of Stakeholder Conflict of Urban Regeneration Sustainability in China: the Application of Social-Based Solutions." ***Sustainability,*** Vol. 11, No. 19, (2019), 5271. https://doi.org/10.3390/su11195271
26. Caputo, A. "Systemic Stakeholders' Management for Real Estate Development Projects" ***Global Business and Management Research: An International Journal,*** Vol. 5, No. 1, (2013), 66-82.
27. Xue, Y. and Xiang, P. "The Social Risk of High-Speed Rail Projects in China: A Bayesian Network Analysis." ***Sustainability,*** Vol. 12, No. 5, (2020), 2087. https://doi.org/10.3390/su12052087
28. Çelik, T., Kamali, S. and Arayici, Y. "Social cost in construction projects." ***Environmental Impact Assessment Review,*** Vol. 64, (2017), 77–86. https://doi.org/10.1016/j.eiar.2017.03.001
29. Chen, Y. Q., Zhang, Y. B. and Zhang, S. J. "Impacts of Different Types of Owner-Contractor Conflict on Cost Performance in Construction Projects". ***Journal of Construction Engineering and Management,*** Vol. 140, No. 6, (2014), 04014017. https://doi.org/10.1061/(ASCE)CO.1943-7862.0000852
30. Mehregana, M. R., Hosseinzadeha, M. and Kazemia, A. "An application of Soft System Methodology." ***Procedia - Social and Behavioral Sciences,*** Vol. 41, (2012), 426 – 433. https://doi.org/10.1016/j.sbspro.2012.04.05
31. Witrianto, "Manajemen Konflik Dalam Pembangunan Masyarakat di Mimika Papua," ***Jurnal Penelitian Sejarah dan***

Budaya, Vol. 1, No. 2, (2015), 283-307. https://doi.org/10.36424/jpsb.v1i2.94

32. Antunes, C. H., Dias, L., Dantas, G., Mathias, J. and Zamboni, L. "An application of Soft Systems Methodology in the evaluation of policies and incentive actions to promote technological innovations in the electricity sector." ***Energy Procedia,*** Vol. 106, (2016), 258 – 278. https://doi.org/10.1016/j.egypro.2016.12.121
33. Magsi, H., Torre, A., Liu, Y. and Sheikh, M. J. "Land Use Conflicts in the Developing Countries: Proximate Driving Forces and Preventive Measures." ***The Pakistan Development Review,*** Vol. 56, No. 1, (2017), 19–30. https://doi.org/10.30541/v56i1pp.19-30
34. Molwus, J. J., Erdogan, B. and Ogunlana, S. "Using structural equation modelling (SEM) to understand the relationships among critical success factors (CSFs) for stakeholder management in construction." Engineering, Construction and Architectural Management, Vol. 24, No. 3, (2017), 426-450. https://doi.org/10.1108/ECAM-10-2015-0161
35. Nguyen, L. H. and Watanabe, T. "The Impact of Project Organizational Culture on the Performance of Construction Projects." ***Sustainability,*** Vol. 9, No. 5, (2017), 781. https://doi.org/10.3390/su9050781
36. Taheri, A. M. J., Haghighi, F. R., Eshtehardian, E. and Abessi, O. "Optimization of Time, Cost and Quality in Critical Chain Method Using Simulated Annealing," ***International Journal of Engineering, Transactions B: Applications***, Vol. 30, No. 5, (2017), 627-635. http://dx.doi.org/10.5829/idosi.ije.2017.30.05b.01

Persian Abstract

چکیده

تعارض اجتماعی در پروژه ها نمایشی سازش ناپذیر از اختلاف در منافع است. هر پروژه بر اساس شرایط آنها از طرف ذینفعان ادعاهای مختلفی دریافت می کند. علاقه پروژه به سطح و اهداف مورد نظر بستگی دارد. متعاقباً ، جامعه محلی پیرامون پروژه منافع آنها را بر اساس انتظارات اقتصادی ، اجتماعی ، فرهنگی و زیست محیطی تغییر می دهد. این مطالعه با هدف آشکار کردن عوامل منافع بین پروژه و جامعه محلی در مورد بروز تعارضات اجتماعی در پروژه ها انجام شده است. از روش سیستم نرم و CATWOE برای شناسایی علت اصلی و عوامل منافع در منازعات اجتماعی استفاده شد. از طریق تحقیقات کتابخانه ای و تحقیقات مرتبط قبلی ، مشخص شد که هشت عامل منافع وجود دارد که باعث ایجاد تضاد اجتماعی در پروژه ها می شود. توصیه می شود که تحقیقات آینده باید از فاکتورهای روش رسمی سیستم (FSM)برای مدل سازی پیش بینی تعارض اجتماعی در پروژه ها بر اساس علایق مختلف بین پروژه و جامعه محلی استفاده کنند. نتایج برای ترکیب محیط استاندارد و چارچوب اجتماعی (ESF)به عنوان یک مفهوم "منافع راه میانه" برای کاهش تعارض اجتماعی در پروژه ها مفید است. با استفاده از ESF صحیح می توان منافع و تعارضات را به خوبی مدیریت کرد تا از مزایای آن برخوردار شده و جامعه و پروژه خود بهره مند شوند. سرانجام ، می توان از نتایج این مطالعه برای تدوین شاخص های درگیری های اجتماعی در پروژه های ساختمانی استفاده کرد.

International Journal of Engineering

Journal Homepage: www.ije.ir

A Study of Blast-induced Vibration on Oil Pipelines based on Numerical and Field Analysis

H. Bakhshandeh Amnieh*, S. M. Mahdi Mirabedi, M. Rahmanpour, V. Jafari

School of Mining, College of Engineering, University of Tehran, Iran

PAPER INFO

Paper history:
Received 27 February 2021
Received in revised form 07 July 2021
Accepted 29 July 2021

Keywords:
Blasting
Ground Vibration
Mining
Numerical Simulation
Oil Piplines
Vibration

ABSTRACT

Blasting is the initial stage of development in mining operations. Therefore, the use of explosives requires a technical design to control its adverse effects on nearby structures. In that regard, the blast vibrations in Izeh-Karun 3 main road project were recorded using four 3-component Blast Recorder seismographs. The seismographs recorded a peak particle velocity of 8.8 mm.s^{-1} in the nearby oil pipe. The blast pattern and the resulting ground vibration are simulated. The numerical model is verified using the recorded seismic data and the empirical model. The stresses applied on the oil pipeline were measured by the static analysis of the stress induced by the oil pipeline's internal pressure and the dynamic analysis of ground vibration. The pipeline stress was equal to 271 MPa, lower than the pipeline yield stress (414 MPa). Therefore, the vibrations induced by the blasting operation did not damage the oil pipeline. Comparing the vibration induced in the oil pipeline (8.8 mm.s^{-1}) with the critical vibration level (50 mm.s^{-1}) showed that the pipelines near the blast operation were at a safe distance.

doi: *10.5829/ije.2021.34.09c.09*

NOMENCLATURE

PPV	Peak Particle Velocity (mm/s)	**Subscripts**	
Q_{max}	Maximum charge per delay (kg)	*Ver*	Vertical component of PPV (mm.s^{-1})
R	Distance from the blast center (m)	*Rad*	Radial component of PPV (mm.s^{-1})
$SD = \frac{R}{\sqrt{Q_{max}}}$	Scaled distance (m.kg$^{-0.5}$)	*Tan*	Tangential component of PPV (mm.s^{-1})

1. INTRODUCTION

Due to the expansion of civil and mining activities, explosive use for rock crushing has increased. The adverse effects of blasting operations include ground vibration, air blast, and fly rock. Ground vibration can cause serious and irreparable damages. In the areas near the blast site, these effects are considered a severe threat to the surface and subsurface structures such as residential and office buildings, bridges, dams, roads, energy carrier, and supply facilities, road tunnels and servicing, underground caves, oil, gas, and water transmission pipelines and walls of mines. Hence, ground vibration as an environmental concern may invoke locals to stand against mining operations. Therefore, it is necessary to measure, predict and control the intensity of ground vibration. Since the beginning of ground vibration studies, different approaches and methods have been applied to evaluate the ground vibration range. The United States Bureau of Mines followed up ground vibration surveys and reviews from 1935 to 1942. As a result of these studies, particle acceleration became a basis for measuring and predicting damage to surface structures. Generally, blast-induced vibrations is expressed based on Peak Particle Velocity (PPV).

Researchers have presented different relations based on seismic data, each of which takes into consideration different parameters. Most of those relations express PPV as a function of the distance from the explosion center and the maximum charge per delay. Relation 1 shows the

*Corresponding Author Institutional Email: hbakhshandeh@ut.ac.ir
(H. Bakhshandeh Amnieh)

Please cite this article as: H. Bakhshandeh Amnieh, S. M. Mahdi Mirabedi, M. Rahmanpour, V. Jafari, A Study of Blast-induced Vibration on Oil Pipelines based on Numerical and Field Analysis, International Journal of Engineering, Transactions C: Aspects, Vol. 34, No. 09, (2021) 2116-2123

general form of empirical relationships in predicting blast-induced ground vibration [1].

$$PPV = K.R^{-B}.Q_{\max}^{A} \quad (1)$$

where K is the empirical constant affected by geological parameters, B and A are the empirical constants affected by the blast design parameters. Q_{max} is the maximum charge per delay (kg), and R is the distance from the blast center (m).

Despite the development of oil, gas, and water transmission pipelines in recent years, very few studies have analyzed pipeline design against blast-induced ground shock wave propagation [2]. Most of the analytical studies performed so far have applied seed waveform overlapping techniques to estimate the blast-induced ground vibration [3]. Kouretzis et al. [2] presented an analytical calculation of surface blast-induced strains to buried pipelines with Three Basic Assumptions of blast-induced vibration on pipelines, Thin-walled, Elastic, and 3D Pipeline Modeling regardless of the slip between soil and pipe. Dagasan et al. [4] conducted tests near the high-pressure gas pipelines to measure the amount of vibration on the pipe and compared them with different criteria and standards.

The ability to do computer calculations has rapidly increased and since the mid-19th century, and the use of these computations has facilitated numerical simulation in many complex processes like blasting [5]. Ma et al. [6] simulated blast waves' propagation in the rock mass with Autodyn and analyzed the explosive, air, and rock interactions. Using ANSYS Autodyn, Park et al. [7] modeled the explosion process in a blast hole through a nonlinear hydrocode. Olarewaju et al. [8] investigated the response of buried pipelines to earthquake-induced vibrations using the FEM numerical method. Xu et al. [9] studied the dynamic response of buried pipelines to the dynamic blast loading through numerical analysis. Jones-Wilkins-Lee equation has been used for explosive behavior modeling. Mitelman and Elmo [10] predicted blast-induced damages to an underground tunnel utilizing a combination of ANSYS Autodyn and ELFEN. Fakhimi and Lanari [11] developed a hybrid of the DEM-SPH numerical model for modeling rock blasting. Applying the FEM numerical model, Yu et al. [12] investigated the impact of blast-induced vibrations on tunnels in soft grounds. Song et al. [13] examined X70 pipeline response under blast loading by a field study and numerical modelling. Jayasinghe et al. [14] studied the blasting effect on the adjacent piles by numerical methods. Abedi et al. [15] developed a mathematical model based on the theory of beam on the elastic foundation to study the behavior of buried pipelines exposed to surface blasting. In addition, they constructed a three-dimensional finite element model, compared the analytical and numerical results, and concluded that the analytically calculated PPVs were higher than the numerical ones. Yung et al. [16] used a three-dimensional numerical model to assess the safety of the DN 1200 buried water pipeline concerning the vibration loads from a tunnel blasting operation. They also evaluated the safety of empty and full pipelines.

To investigate the effects of surface blasting on Monon pipelines and the dynamic response of the buried pipelines to it; Tang et al. [17] conducted dynamic vibration studies and dynamic strain monitoring in Zhushan with an area of around 100×600 sqm. They studied the peak particle velocity of pipelines under different blast charges at different distances. Finally, they presented a relationship between the PPV of the ground and the pipeline. Wang et al. [18] experimentally and numerically investigated the effect of blasting-induced vibrations on gas pipelines. They analyzed dynamic parameters such as PPV and PPA and proposed equations to predict blast vibration parameters based on the distance. Numerical modeling and seismic analysis are also applicable in the design of oil storage facilities [19]. Hassani and Basirat [19] investigate the subsidence effect on buried pipelines through numerical modeling. Apart from numerical studies, some researchers used statistical and soft computing techniques to predict PPV [20-26]. The main drawback of these models is that they need excessive data and new studies in this field requires a collection of expanded database and test of new models. Moreover, in engineering project where the safety of infrastructures (such as buried pipelines) is of more interests, application of prediction models should be chosen carefully. While, numerical models enforced by field measurements will provide consistent result which is the motivation of current paper. In that regard, a road construction project near the vicinity to some oil and gas pipelines is considered for evaluation.

Karun 3 main road to Izeh has a length of 21 km. Due to its vicinity to the oil and gas pipelines and the possibility of damage imposed to the pipelines by the nearby blasting operation and the resulted blast-induced ground vibration, a careful controlled blasting operation is necessary. Hence a detailed seismic study is carried out to confirm the blasting parameters. In that regard, numerical modelling is used to model the blast-induced ground vibration and study its damages to the nearby oil and gas pipelines. Numerical analysis of blast-induced ground vibration is consist of two parts, (1) modeling the blast process and (2) simulating vibration wave propagation near the oil pipelines. Therefore, this paper adopts a blasting experiment near an on-site buried pipeline, while ground vibrations are monitored. These measurements will verify the numerical model. Field measurements reflect the actual behavior of the ground and the pipelines. Finally, the model is used to study the pipeline response to the nearby surface blast load.

2. MATERIAL AND METHODS

2. 1. Field Measurements Izeh-Karun's main road project is crossing the Sarvak-Ilam formations. These formations are almost limestone. The road project is close to existing oil and gas pipelines. Due to its vicinity to the pipelines there is a possibility of any damage to those pipelines. These damages may be a result of blast-induced ground vibration. Hence, controlled blasting and detailed seismic studies should be carried out. In that regard, a blasting experiment near an on-site buried pipeline is conducted and ground vibrations are recorded. For this purpose, four 3-component Blast Recorder Seismographs (BRS) were used to record blast-induced ground vibration. These BRSs are installed along the blasting block and the pipelines. The relative position of the oil and gas pipelines and the blasting block are shown in Fig. 1. This figure shows the location of seismic record stations where a BRS is installed. According to this figure, St2, St3, and St4 are located in a particular direction between the blasting block and the pipelines. St2 is 46.6 meters, and St4 is 87.5 meters away from the blasting block.

As stated, a blasting experiment near an on-site buried pipeline is conducted. The blasting block contains 12 blastholes with a depth of 3 meters. The blastholes diameter is 3 inches, and they are drilled in 2 rows. Burden and spacing are 2 and 2.5 meters, respectively. 38% of the blasthole (i.e., 1.15 meters) is charged with ANFO, while the rest of the blastholes (i.e., 1.85 meters) are used for stemming. In this operation, the NONEL system with 42-millisecond connectors is used. The initiating pattern of blastholes is given in Figure 2a.

2. 2. Numerical Simulation of Blasting Wave Propagation Autodesk inventor software package has been used to create the model geometry. Figure 2b shows the geometry of the model and the blasting block location. The model dimensions are 120, 20, and 32 meters in length, width, and height. The mesh elements must be within the range of 0.1 to 0.125 of the propagated wavelengths to capture wave propagation accurately. Large mesh dimensions for long wavelengths can result in reasonable accuracy, but it doesn't have good accuracy for shorter wavelengths [27, 28].

For this reason, mesh generation follows a dimension of 30 cm. In dynamic numerical modeling of waves propagating, the model boundaries must be determined so that the elastic waves are allowed to exit the model. This strategy will eliminate the reflection errors within the model [29]. The impedance of model boundaries are assumed as transmitting boundary that provides a condition which prevent wave reflections inside the model.

2. 2. 1. Rock Mass Characteristics The project is located on Sarvak-Ilam formation, and the dominant rock type in the area is limestone. Since the pipeline is far enough from the blasting block then it is located in the elastic zone. The model environment is assumed as continuous media. The Drucker-Prager model was considered as the constitutive model of limestone. The behavioral parameters of the limestone are shown in Table 1.

2. 2. 2. Explosive Characteristics The main explosive used in the project was ANFO. The Jones-Wilkins-Lee (JWL) model describes the ANFO behavioral equation (Equation (2)) [30].

$$P = A\left(1-\frac{\omega}{R_1V}\right)e^{-R_1V} + B\left(1-\frac{\omega}{R_2V}\right)e^{-R_2V} + \frac{\omega E}{V} \quad (2)$$

where, A, B, C, R1, R2 and ω are, constants. Table 2 summarizes the parameters and their values for ANFO.

2. 2. 3. Characteristics of the Pipeline and the Surrounding Soil In the execution of oil pipelines, soft soil cushions with 5 meters distance from each other

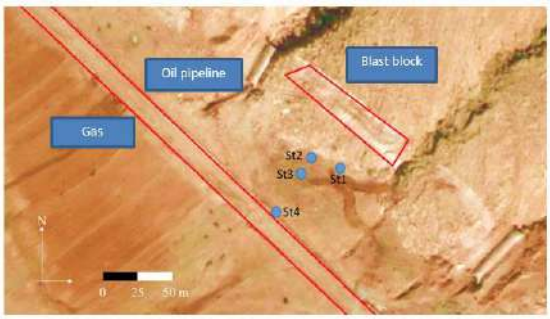
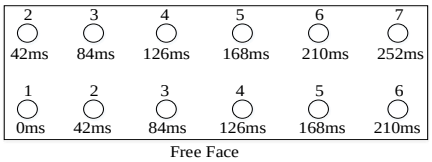
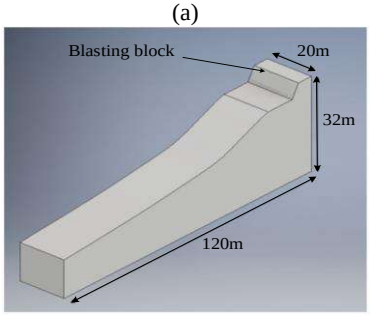


Figure 1. Location of oil and gas pipelines, seismic stations and the blasting block



(a)



(b)

Figure 1. a) Blasting pattern and their timing, b) 3D geometric modeling

are usually placed at the channel bottom before laying the pipelines. Then the pipelines are covered with soft sandy soil. The constitutive models considered for the oil pipeline and its surrounding soil are linear Drucker-Prager and linear elastic. The resistance parameters and specifications of the pipeline are stated in Table 3. Moreover, the filling soil specifications around the pipeline are given in Table 4.

The 3D model of the pipeline is shown in Figure 3. It should be said that the shell element is used to discretize the pipeline. And the solid element is used to discretize the surrounding material. The pipeline's static analysis showed that the appropriate dimension of the pipeline's mesh was 6 cm. Concerning the resolution of dynamic conditions, a mesh dimension of 30 cm is selected for the pipeline's surrounding environment.

TABLE 1. Dynamic parameters of limestone

Dynamic features of rock mass	Unit	Quantity
Density	$kg.m^{-3}$	2640
Modulus of elasticity	GPa	14.46
Poisson's ratio	-	0.44
Bulk module	GPa	40.17
Shear module	GPa	5.02
Uniaxial compressive strength	MPa	65
Uniaxial tensile strength	MPa	5.5

TABLE 2. JWL parameters and their values for ANFO

Parameter	Unit	Quantity
Density	$Kg.m^{-3}$	850
Velocity of Detonation	$m.s^{-1}$	4160
Weight Unit Energy	$MJ.kg^{-1}$	2.668
E	GPa	5.15
A	-	49.46
B	-	1.891
R1	-	3.907
R2	-	1.118
ω	-	0.333

TABLE 3. Pipeline Specifications

Parameter	Unit	Quantity
Material	-	API-5LX60
Operational pressure	MPa	5.2
Thickness	mm	7.14
Diameter	mm (in)	762 (30)
Stress yields	MPa (psi)	414 (60000)
Elastic module	GPa	207
Poisson's ratio	-	0.3
Density	$Kg.m^{-3}$	7850

TABLE 4. Pipeline surrounding Specifications [35]

Parameter	Unit	Quantity
Density	$Kg.m^{-3}$	1750
Elastic module	MPa	100
Poisson's ratio	-	0.35
Internal friction angle	degree	30

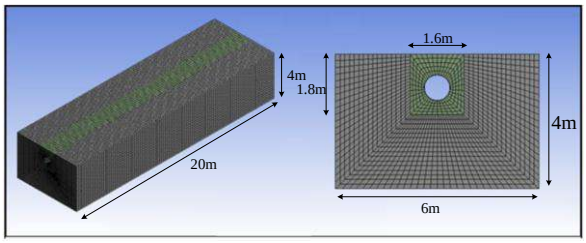


Figure 3. The pipeline and surrounding environment meshing

3. RESULTS AND DISCUSSION

The recorded seismic data was examined to determine PPV. The recorded data are processed in Origin software. In this process, instrument effect are removed and blasting block section was separated from the other parts by binary conversion. Finally, the PPV components are obtained in every station, and the results are given in Table 6. The PPV diagram is compared to the maximum charge per delay as shown in Figure 4. Checking the seismic data showed that the blast-induced ground vibration imposed to the oil pipeline was equal to 8.8 mm/s. According to Figure 1, the seismograph stations St2, St3, and St4 are located between the blasting block and the pipeline. Therefore, these three stations are considered for comparison in the rest of the paper.

After analyzing particle displacement velocity diagrams, PPV was determined separately at the stations. The numerical model is solved, and the resulting PPV was compared with field measurements. The comparison is made in each measurement station for the three tangential, vertical, and radial components of PPV. Since PPV decreases over time, and in order to reduce the computational time, problem-solving time is considered 1000 ms. Figure 5 shows the particle velocity chart for the station St4. In table 7, the predicted PPV is stated for all the radial, vertical and tangential components.

TABLE 6. The seismic data related to the blast operation

Seismograph station	St1	St2	St3	St4
R (m)	48.1	46.6	58.5	87.5
SD (m.kg-0.5)	16.6	16.1	20.2	30.2
PPV_{Ver} (mm.s^{-1})	6.9	12	7.8	4.9
PPV_{Ver} (mm.s^{-1})	10.6	16.3	15.7	6.5
PPV_{Ver} (mm.s^{-1})	10.6	11.7	6.3	3.3
Resulting PPV (mm.s^{-1})	16.5	23.4	18.6	8.8

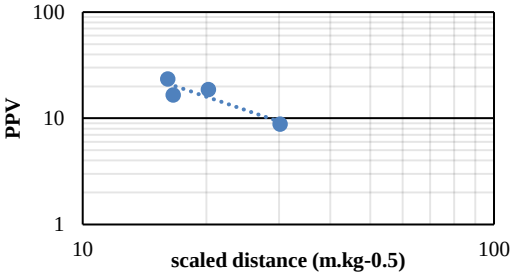


Figure 4. The PPV compared to the scaled distance

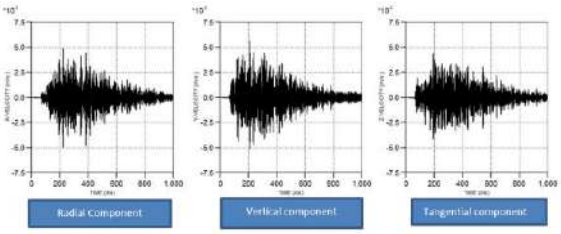


Figure 5. Particle velocity with time for radial, vertical and tangential components of St4 (R=87.5 m)

3. 1. Validation of Numerical Modeling with Seismic Data Figure 6 shows the PPV graph concerning the scaled distance (SD) for the three radial, vertical, and tangential components. The average error of 10.4% between numerical analysis and field measurement shows a good match between numerical simulation and recorded field data.

3. 2. Validation of Numerical Modeling with Empirical Models Many researchers have discussed empirical relations for predicting ground vibration. Most of these models express PPV based on the scaled distance. By studying the data from 659 seismographs in limestone, Ozer [31] presented the empirical model in Equation 3. Considering the similarity of the conditions with Ozer's data, this model was used to recheck the numerical modeling.

$$PPV = 3019.95\ SD^{-1.69} \quad (3)$$

In order to compare the numerical results and the empirical values calculated by Equation (3), the PPV vs. scaled distance was drawn (Figure 7). Moreover, in Table 8, the model error was calculated by comparing the results. According to Table 9, the average error of the numerical modeling results and Ozer's empirical model is 9.1%, indicating reasonable numerical modeling accuracy.

TABLE 7. Predicted PPV at different stations by the model

Station	R (m)	SD (m.kg-0.5)	PPV_{Ver}	PPV_{Ver}	PPV_{Ver}
St2	46.6	16.1	13.2	13	11.6
St3	58.5	20.2	7.1	14.3	7.5
St4	87.5	30.2	4.92	6.9	4.3

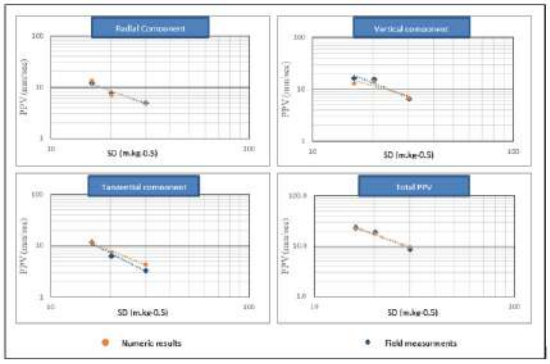


Figure 6. PPV diagram recorded by seismographs and numerical modeling with scaled distance for PPV components

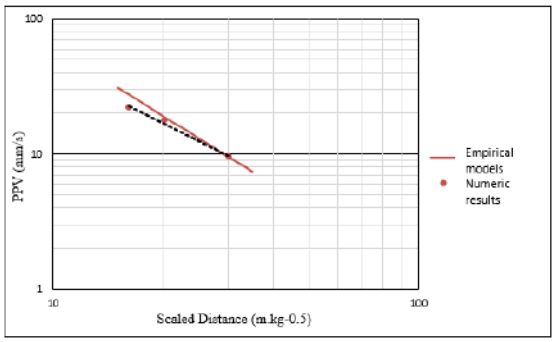


Figure 7. Comparing the numerical results and Ozer's model

TABLE 8. Comparison of the numerical and empirical results

Measurement point	PPV (mm/s)		error (%)
	Numerical model	Empirical model	
St2	21.9	27.6	20.7
St3	17.6	18.8	6.3
St4	9.5	9.52	0.2

3. 3. Blast Induced Ground Vibration Effects on Pipelines First, a separate static analysis is conducted. In static analysis, oil pressure inside the pipeline, displacements, and the induced stresses was measured. Then, ground vibration near the oil pipeline was investigated by dynamic analysis. In this study, the fluid weight inside the pipeline was neglected. Finally, the blast-induced ground vibration in the pipelines' vicinity was compared with different standards.

After the model's equilibrium, the liquid's operational pressure inside the pipe and the initial condition of dynamic analysis are determined. The results showed that the fluid's maximum stress was equivalent to 255 MPa and the maximum displacement in the upper part of the pipe was 821 µm. The ground vibration was simulated as a dynamic load applied to the model in 1100 ms. Figure 8 shows the vibration diagram applied to the pipeline. Considering pipeline specifications, the Von-Mises equivalent stress is 274.9 MPa. It indicates the accuracy of the static analysis compared to the equivalent stress obtained by numerical modeling (i.e., 255 MPa). Then the vibration velocity and Von-Mises equivalent stress for the pipeline were obtained. The maximum equivalent stress on the pipeline was 271 MPa. Figure 9 shows the stress diagram of the Von-Mises equivalent stress at the most critical point of the pipeline. Considering the yield stress of the oil pipeline (i.e., 414 MPa), the equivalent stress caused by the internal pressure of the pipe and the blast-induced vibration is less than the yield stress. It means that the blasting operation may not damage the oil pipeline. Moreover, the vulnerability of the pipelines near the blast operation was measured by comparing the effect of the blast-induced ground vibration on the oil and gas pipelines with several standards (Table 9). Considering that the most critical PPV is for Canadian Standard, which is equal to 50 mm/s, and the PPV applied to the oil pipeline is 8.8 mm/s, one could conclude that the vibration in the oil pipeline is lower than all

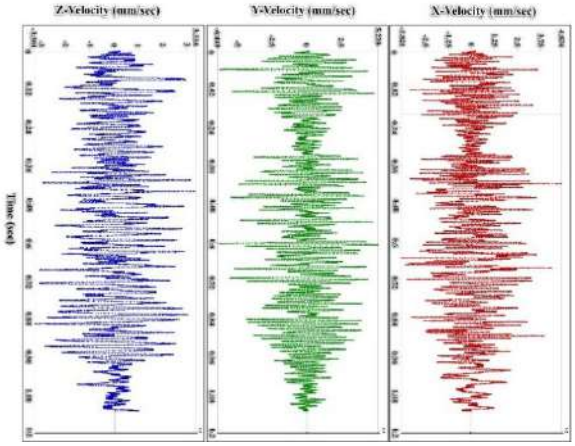


Figure 8. The pipeline vibration in radial, vertical and tangential components

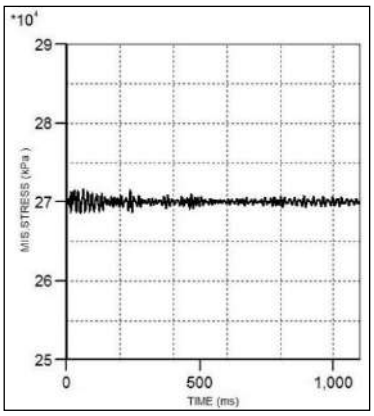


Figure 9. Equivalent stress applied to the pipeline

TABLE 9. Allowable PPV for gas and pipelines in different standards

Standards	PPV ($mm.s^{-1}$)
United State Bureau of Mines	**127**
DIN_4150-3	**100**
Northern Gas Network	**75**
Canada	**50**
API	**50.8**
Austria Standard	**75**

permissible limits. Thus the blasting operation does not damage the pipelines.

4. CONCLUSION

The study of controlled blasting and detailed seismic analysis is necessary for safety analysis of infrastructure, which is the motivation of this paper. In this paper combining the field measurements, the safety evaluation of the buried pipeline under surface blast loads is studied. In the engineering project with similar conditions, the buried pipeline will operate normally. In that regard, numerical modeling is used to study the blast-induced ground vibration and evaluate its damages to the nearby oil and gas pipelines. Some field measurements verify the numerical model before any further applications. Through numerical modeling of a blast process, the propagation of elastic waves in the rock mass and the probability of any damage to the oil and gas pipelines were studied, and the following results were obtained.

After recording and analyzing seismograph data, a PPV of 8.8 $mm.s^{-1}$ at the seismic station near the pipeline was recorded. Model verification results show a good relationship with the propagation of vibration waves based on the seismic data and the empirical relation. The results' average error compared with seismic data and the

Ozer empirical model is equal to 10.4% and 9.9%, respectively.

By dynamic analysis of the fluid pressure inside the pipeline and the ground vibration applied to the pipeline, the Von-Mises equivalent stress was measured to be 271 MPa, which is almost 35% less than the yield stress of the oil pipeline. As a result, the blast-induced vibrations do not damage the oil pipeline. According to the standards, the critical threshold for oil pipelines is 50 $mm.s^{-1}$. Analysis of the seismic data and its comparison to that standard indicated that the maximum vibration applied to the oil pipeline (8.8 $mm.s^{-1}$) is acceptable. Therefore, according to the main roads' vibration standards, the oil and gas pipelines near the Karun 3 road to Izeh are safe.

5. REFERENCES

1. Rai, R. and Singh, T., "A new predictor for ground vibration prediction and its comparison with other predictors", (2004).
2. Kouretzis, G.P., Bouckovalas, G.D. and Gantes, C.J., "Analytical calculation of blast-induced strains to buried pipelines", ***International Journal of Impact Engineering***, Vol. 34, No. 10, (2007), 1683-1704.
3. Blair, D., "Non-linear superposition models of blast vibration", ***International Journal of Rock Mechanics and Mining Sciences***, Vol. 45, No. 2, (2008), 235-247.
4. Dağaşan, Y., Erten, O. and Bilgin, H., "The analysis of ground vibrations induced by foundation excavation test blasts nearby a buried high pressure gas pipeline", in 8th Drilling-Blasting Symposium. (2015).
5. Bobet, A., Fakhimi, A., Johnson, S., Morris, J., Tonon, F. and Yeung, M.R., "Numerical models in discontinuous media: Review of advances for rock mechanics applications", ***Journal of Geotechnical and Geoenvironmental Engineering***, Vol. 135, No. 11, (2009), 1547-1561, doi: 10.1061/(ASCE)GT.1943-5606.0000133.
6. Ma, G., Hao, H. and Zhou, Y., "Modeling of wave propagation induced by underground explosion", ***Computers and Geotechnics***, Vol. 22, No. 3-4, (1998), 283-303, doi: 10.1016/S0266-352X(98)00011-1.
7. Park, D., Jeon, B. and Jeon, S., "A numerical study on the screening of blast-induced waves for reducing ground vibration", ***Rock Mechanics and Rock Engineering***, Vol. 42, No. 3, (2009), 449-473, doi: 10.1007/s00603-008-0016-y.
8. Olarewaju, A., Rao, N. and Mannan, M., "Response of underground pipes due to blast loads by simulation- an overview", ***Electronic Journal of Geotechnical Engineering***, Vol. 15, No. H, (2010), 831-852.
9. Xu, G.F., Deng, Z.D., Deng, F.F. and Liu, G.B., "Numerical simulation on the dynamic response of buried pipelines subjected to blast loads", in Advanced Materials Research, Trans Tech Publ. Vol. 671, (2013), 519-522.
10. Mitelman, A. and Elmo, D., "Modelling of blast-induced damage in tunnels using a hybrid finite-discrete numerical approach", ***Journal of Rock Mechanics and Geotechnical Engineering***, Vol. 6, No. 6, (2014), 565-573, doi: 10.1016/j.jrmge.2014.09.002.
11. Fakhimi, A. and Lanari, M., "Dem–sph simulation of rock blasting", ***Computers and Geotechnics***, Vol. 55, (2014), 158-164, doi: 10.1016/j.compgeo.2013.08.008.
12. Yu, H., Yuan, Y., Yu, G. and Liu, X., "Evaluation of influence of vibrations generated by blasting construction on an existing tunnel in soft soils", ***Tunnelling and Underground Space Technology***, Vol. 43, (2014), 59-66, doi: 10.1016/j.tust.2014.04.005.
13. Song, K., Long, Y., Ji, C., Gao, F. and Chen, H., "Experimental and numerical studies on the deformation and tearing of x70 pipelines subjected to localized blast loading", ***Thin-Walled Structures***, Vol. 107, (2016), 156-168, doi: 10.1016/j.tws.2016.03.010.
14. Jayasinghe, L.B., Zhou, H., Goh, A., Zhao, Z. and Gui, Y., "Pile response subjected to rock blasting induced ground vibration near soil-rock interface", ***Computers and Geotechnics***, Vol. 82, (2017), 1-15, doi: 10.1016/j.compgeo.2016.09.015.
15. Abedi, A.S., Hataf, N., Shivaei, S. and Ghahramani, A., "Comparative study of analytical and numerical evaluation of the dynamic response of buried pipelines to road-cut excavation blasting", ***Geomechanics and Geoengineering***, Vol. 15, No. 2, (2020), 140-148, doi: 10.1080/17486025.2019.1634289.
16. Xia, Y., Jiang, N., Zhou, C. and Luo, X., "Safety assessment of upper water pipeline under the blasting vibration induced by subway tunnel excavation", ***Engineering Failure Analysis***, Vol. 104, (2019), 626-642, doi: 10.1016/j.engfailanal.2019.06.047.
17. Tang, Q., Jiang, N., Yao, Y., Zhou, C. and Wu, T., "Experimental investigation on response characteristics of buried pipelines under surface explosion load", ***International Journal of Pressure Vessels and Piping***, Vol. 183, (2020), 104101, doi: 10.1016/j.ijpvp.2020.104101.
18. Wang, K., Liu, Z., Qian, X. and He, Y., "Dynamic characteristics and damage recognition of blast-induced ground vibration for natural gas transmission pipeline and its integrated systems", ***Mechanical Systems and Signal Processing***, Vol. 136, (2020), 106472, doi: 10.1016/j.ymssp.2019.106472.
19. Hassani, R. and Basirat, R., "The investigation of subsidence effect on buried pipes in 3d space", ***International Journal of Engineering***, Vol. 30, No. 8, (2017), 1182-1189, doi: 10.5829/ije.2017.30.08b.10.
20. Krishnamoorthy, A., "Finite element method of analysis for liquid storage tank isolated with friction pendulum system", ***Journal of Earthquake Engineering***, Vol. 25, No. 1, (2021), 82-92, doi: 10.1080/13632469.2018.1498815.
21. Khandelwal, M. and Singh, T., "Prediction of blast-induced ground vibration using artificial neural network", ***International Journal of Rock Mechanics and Mining Sciences***, Vol. 46, No. 7, (2009), 1214-1222, doi: 10.1016/j.ijrmms.2009.03.004.
22. Amiri, M., Amnieh, H.B., Hasanipanah, M. and Khanli, L.M., "A new combination of artificial neural network and k-nearest neighbors models to predict blast-induced ground vibration and air-overpressure", ***Engineering with Computers***, Vol. 32, No. 4, (2016), 631-644, doi: 10.1007/s00366-016-0442-5.
23. Hasanipanah, M., Faradonbeh, R.S., Amnieh, H.B., Armaghani, D.J. and Monjezi, M., "Forecasting blast-induced ground vibration developing a cart model", ***Engineering with Computers***, Vol. 33, No. 2, (2017), 307-316, doi: 10.1007/s00366-016-0475-9.
24. Fouladgar, N., Hasanipanah, M. and Amnieh, H.B., "Application of cuckoo search algorithm to estimate peak particle velocity in mine blasting", ***Engineering with Computers***, Vol. 33, No. 2, (2017), 181-189, doi: 10.1007/s00366-016-0463-0.
25. Yang, H., Rad, H.N., Hasanipanah, M., Amnieh, H.B. and Nekouie, A., "Prediction of vibration velocity generated in mine blasting using support vector regression improved by optimization algorithms", ***Natural Resources Research***, Vol. 29, No. 2, (2020), 807-830.
26. Azimi, Y., "Prediction of seismic wave intensity generated by bench blasting using intelligence committee machines",

International Journal of Engineering, Vol. 32, No. 4, (2019), 617-627, doi: 10.5829/ije.2019.32.04a.21.

27. Armaghani, D.J., Hasanipanah, M., Amnieh, H.B., Bui, D.T., Mehrabi, P. and Khorami, M., "Development of a novel hybrid intelligent model for solving engineering problems using gs-gmdh algorithm", ***Engineering with Computers***, (2019), 1-13, doi: 10.1007/s00366-019-00769-2.
28. Hibbitt, K., & Sorensen., "Abaqus/cae user's manual, Hibbitt, Karlsson & Sorensen Incorporated, (2002).
29. J Swanson, J.A., "Ansys autodyn user manual, V18.2. Canonsburg, (2018).
30. Davis, L.L. and Hill, L.G., "Anfo cylinder tests", in AIP conference proceedings, American Institute of Physics. Vol. 620, No. 1, (2002), 165-168.
31. Ozer, U., "Environmental impacts of ground vibration induced by blasting at different rock units on the kadikoy–kartal metro tunnel", ***Engineering Geology***, Vol. 100, No. 1-2, (2008), 82-90.

Persian Abstract

چکیده

عملیات انفجار به‌عنوان یکی از مراحل اصلی چرخه عملیات معدنی است. در استفاده از مواد منفجره باید تاثیر منفی فرایند انفجار بر سازه های اطراف کنترل شوند و این امر نیازمند طراحی فنی است. برای بررسی تأثیر لرزش‌های ناشی از انفجار بر خط لوله نفت، لرزش‌های ایجادشده از عملیات انفجار پروژه راه اصلی ایذه–کارون۳ توسط ۴ لرزه‌نگار سه مؤلفه‌ای ثبت شد. بر اساس نتایج میدانی، حداکثر سرعت ذرات در مجاورت خط لوله نفت ۸/۸ میلیمتر بر ثانیه ثبت شد. پس از ثبت داده‌ها، الگوی انفجار و لرزش‌های ناشی از انفجار مدلسازی شدند و مدل بر اساس داده‌های ثبت شده و همچنین مدل‌های تجربی اعتبارسنجی شد. سپس تنش‌های وارد بر خط لوله محاسبه شدند. این تنش‌ها ناشی از سیال داخل لوله و لرزش زمین هستند. طبق نتایج مدل، مقدار تنش وارد بر خط لوله ۲۷۱ مگاپاسکال است که کمتر از مقاومت لوله (۴۱۴ مگاپاسکال) است. طبق نتایج مدلسازی، لرزش‌های ناشی از انفجار آسیبی به خط لوله نفت وارد نمی‌کندو این لرزش‌ها به طور قابل توجهی کمتر از حد مجاز (۵۰ میلیمتر بر ثانیه) است و در محدوده امن قرار دارند.

International Journal of Engineering

Journal Homepage: www.ije.ir

Energy-aware Task Scheduling in Cloud Compting Based on Discrete Pathfinder Algorithm

A. Zandvakili, N. Mansouri*, M. M. Javidi

Department of Computer Science, Shahid Bahonar University of kerman, Kerman, Iran

PAPER INFO

Paper history:
Received 10 April 2020
Received in revised form 29 July 2021
Accepted 30 July 2021

Keywords:
Resource Utilization
Energy Efficiency
Cloud Computing
Task Scheduling
Throughput
Makespan
Latency

ABSTRACT

Task scheduling is one of the fundamental issues that attract the attention of lots of researchers to enhance cloud performance and consumer satisfaction. Task scheduling is an NP (non-deterministic polynomial)-hard problem that is challenging due to the several conflicting objectives of users and service providers. Therefore, meta-heuristic algorithms are the more preferred option for solving scheduling problems in a reasonable time. Although many task scheduling algorithms are proposed, existing strategies mainly focus on minimizing makespan or energy consumption while ignoring other performance factors. In this paper, we propose a new task scheduling algorithm based on the Discrete Pathfinder Algorithm (DPFA) that is inspired by the collective movement of the animal group and mimics the guidance hierarchy of swarms to find hunt. The proposed scheduler considers five objectives (i.e., makespan, energy consumption, throughput, tardiness, and resource utilization) as cost functions. Finally, different algorithms such as Firefly Algorithm (FA), Particle Swarm Optimization (PSO), Grasshopper Optimization Algorithm (GOA), and Bat Algorithm (BA), are used for comparison. The experimental results indicate that the proposed scheduling algorithm with FA, BA, PSO, and GOA improved up to 9.16%, 38.44%, 3.59%, and 3.44%, respectively. Moreover, the results show dramatic improvements in terms of resource utilization, throughput, and energy consumption.

doi: *10.5829/ije.2021.34.09c.10*

1. INTRODUCTION

1. 1. Cloud Computing Cloud Computing refers to both the applications delivered as services over the internet and the hardware and system software in the datacenters that provide those services [1]. As shown in Figure 1, the benefits of cloud computing are: on-demand self-service, multi-tenancy, offers resilient computing, fast and effective virtualization, offers advanced online security, location, and device independence, always available, and scales automatically to adjust to allow pay-per-use. Cloud computing delivers different services as a utility to users through the internet. One consequence of this model is that large cloud data centers consume large amounts of energy and produce significant carbon footprints. Researchers have recently taken energy consumption seriously as a contribution to expand the green cloud space. In datacenters, many efficient technologies including dynamic voltage and frequency technology, resource hibernation, and memory optimization are utilized for reducing energy consumption. Achieving a reasonable trade-off among energy consumption, resource utilization, and quality of service (QoS) requirements is a challenging problem, especially with diverse tasks in a heterogeneous environment. A common objective of cloud providers is to develop resource provisioning and management solutions that minimize energy consumption [2]. To achieve this objective, a thorough understanding of energy consumption patterns in complex cloud systems is imperative. Resource utilization is the next challenge of cloud computing. Effective resource scheduling not only reduces resource consumption (increases the resource utilization ratio) but also executes incoming tasks in minimum time (minimizes the makespan).

*Corresponding Author Email: najme.mansouri@gmail.com, n.mansouri@uk.ac.ir (N. Mansouri)

Please cite this article as: A. Zandvakili, N. Mansouri, M. M. Javidi, Energy-aware Task Scheduling in Cloud Compting Based on Discrete Pathfinder Algorithm, International Journal of Engineering, Transactions C: Aspects, Vol. 34, No. 09, (2021) 2124-2136

Figure 1. Advantages and disadvantages of cloud computing

At cloud datacenters, inefficient task scheduling may reduce revenue as a result of resource underutilization. In this context, to perform efficient scheduling of tasks on the cloud, the makespan needs to be reduced. In cloud computing, the tardiness parameter is another important objective function. Naturally, a low value for this parameter, makes the response time shorter and thus increases the satisfaction of cloud users. Throughput is the next challenge of cloud computing. Throughput refers to the performance of tasks by a server or device over a specific period. For transaction processing systems, it is normally measured as transactions-per-second. For systems processing bulk data, such as cloud servers, it is measured as a data rate (e.g., Megabytes per second) [3]. Adequate throughput is important to ensure all applications run with optimal efficiency. All these challenges, along with the timely receipt of services with the least delay by the server, depending on the optimal scheduling.

1. 2. Task Sheduling Task scheduling in a distributed heterogeneous computing environment can be identified as a non-linear, multi-objective, NP-hard optimization problem that strives to optimize cloud resource utilization and satisfy the QoS requirements. [4]. Task scheduling is a major challenge in cloud computing, which refers to the technique of mapping a set of tasks to a set of machines to fulfill users' demands. The task scheduling problem can be modeled in different modes. The difference can be based on the number of machines (single or multiple machines), the type of machines (the machines are the same or different), the dependence or independence between tasks. In this paper, machines are capable to perform all tasks, but each task is considered individually (each task does not include smaller parts). Undoubtedly, solving the task scheduling problem is very time-consuming and has a high computational load due to the nature of these types of problems with precise methods, so an acceptable solution can be obtained by choosing the proper meta-heuristic algorithms. There are various types of scheduling algorithms, some of them are static

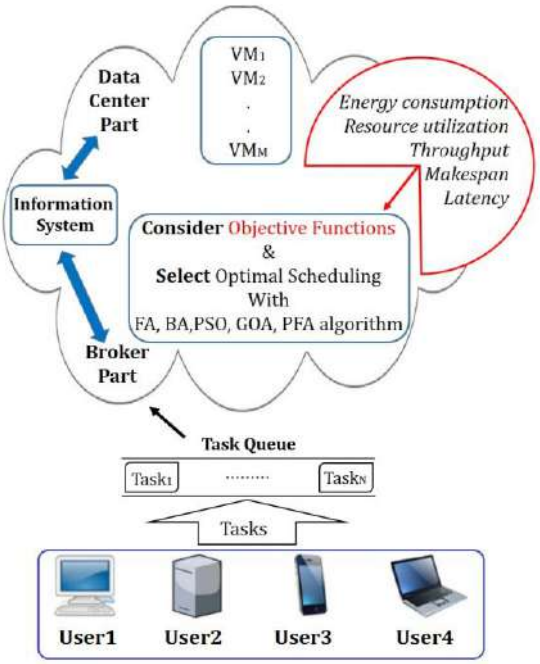


Figure 2. Architecture of task scheduling system

scheduling algorithms that are considered suitable for small or medium scale cloud computing, and dynamic scheduling algorithms that are considered suitable for large scale cloud computing environments.

Figure 2 shows the scheduling system architecture that users send their requests to the cloud and wait for the results. In the cloud environment, the scheduler must select the appropriate machine based on the indicators considered in the objective function and send the tasks to the machines.

Unworthy assignment of tasks to a cloud server can increase the waiting time, makespan, and energy consumption, with inefficient load distribution. An efficient algorithm needs to assign tasks to a suitable set of machines to achieve the desired objectives. Most of the existing task scheduling strategies ignore multiple-objective issues such as energy consumption, makespan, tardiness, throughput, and resource utilization simultaneously, and their main focus is to minimize the cost or completion time of the task scheduling without regarding the QoS metrics. Therefore, the objectives of this paper are to minimize resource utilization, makespan, and energy consumption while maximizing throughput of the cloud servers. We use a meta-heuristic algorithm, named PFA and proposed a new task scheduling algorithm based on Discrete PFA (DPFA). We select PFA for solving task scheduling problem since:

- It can mimic the collective movements of swarms by using the hierarchy between the head and other members of the swarm.

- It has two separate mathematical formulations for position updating of head and other members and tends to move locally in the final steps.
- Can explore the promising solutions, provide the abrupt changes in the initial iterations, and exploit the best one throughout iterations.

In recent years, many optimization algorithms have been used to solve multi-objective task scheduling problem, among them the FA [5], PSO [6], BA [7], and GOA [8] are popular.

Nicolas et al. [9] presented a Multi-Objective Discrete Firefly Algorithm (MO-DFFA) for solving the Flexible Job-shop Scheduling Problem (FJSP). The authors minimized three different objectives simultaneously, the sum of the completion times of the orders, the workload of the critical machine, and the total workload of all machines are their objective functions. They used Genetic Algorithm (GA) and the Greedy Randomized Adaptive Search Procedure (GRASP) for comparison.

The main goal of Kumar and Sharma [10] was to design and develop a task processing framework that has the decision-making capability to select the optimal resource at runtime to process the applications (diverse and complex nature) at virtual machines using modified PSO. The authors proposed algorithm that gives a non-dominance set of optimal solutions and improves various influential parameters (i.e., time, cost, throughput, task acceptance ratio). They used Cloudsim tool environment [11] and PSO, adaptive PSO, artificial bee colony (ABC), BA, and improved min–min load-balancing algorithm for comparison.

Shareh et al. [12] investigated the tasks scheduling problem in open shops using the BA based on ColReuse and substitution meta-heuristic functions. They performed simulations in the MATLAB and used GA for comparison.

In this paper, our main research contributions are as follows:

- Use PFA for scheduling in the cloud computing environment.
- Design a multi-objective scheduling strategy for finding optimal scheduling based on multiple conflicting objectives, namely energy consumption, makespan, tardiness, throughput, and resource utilization.
- Experimental analysis is performed to compare the proposed algorithm with FA, BA, PSO, and GOA.

The rest of the paper is organized as follows: Section 2 discusses the related works, which deal with existing task scheduling techniques in the cloud environment. Mathematical models and problem formulation are discussed in section 3. The technical solution is discussed in section 4. Section 5 deals with performance evaluation and section 6 consists of the conclusion and future work.

2. RELATED WORK

Scheduling is the art of analyzing the required QoS parameters to determine which activity should be performed. There are several conflicting parameters in the task scheduling problem. Using these parameters as objective functions in the optimization algorithms is difficult, time-consuming, and costly (Especially using more than three objective functions). Due to this reason, researchers usees meta-heuristic algorithms to achieve optimal scheduling. Meta-heuristic algorithms have revealed significant performance based on different scheduling approaches.

Agarwal and Srivastava [13] used PSO to minimize the execution time. The proposed PSO-based task scheduling mechanism keeps the overall response time minimum and uses the CloudSim simulator with the existing greedy and GA-based task scheduling mechanism.

Raju and Devarakonda [14] tried to reduce the makespan time in a cloud environment by introducing the new method Modified Greedy PSO with Clustered approach (MGPSOC). The MGPSOC algorithm makes use of clustering with bio-inspired techniques. The authors used PSO and Greedy-PSO algorithms for comparison.

Avinashi et al. [15] presented a novel hybrid method by combining Grey Wolf Optimization (GWO) and PSO. The proposed method optimizes the makespan, execution time, and response time. PSO improves the optimization performance of GWO in the proposed method. The proposed algorithm performance is evaluated with GA and GWO algorithms.

Tabrizchi et al. [16] proposed a novel self-adaptive hybrid Imperialist Competitive Algorithm (ICA)-PSO algorithm for dealing with associate multi-task scheduling problems. The authors combined ICA and PSO to improve the exploration. They used PSO and ICA algorithms for comparison.

Koneti et al. [17] presented a Cost-Effective Firefly-based Algorithm (CEFA) to solve workflow scheduling problems. The proposed CEFA uses a novel method for problem encoding, population initialization, and fitness evaluation intending to provide cost-effective and optimized workflow execution within the limited time. The performance of the proposed algorithm is compared with the PSO algorithm, Robustness-Cost-Time (RCT), Robustness-Time-Cost (RTC), and Regressive Whale Optimization (RWO), in terms of response time and makespan.

The number of iterations is very important in the meta-heuristic algorithm. In the existing approach, the number of iterations is very large which increases the total execution cost and time. Kaur and Mann [18], proposed a Hybrid Cost-Effective Genetic and Firefly Algorithm (GAFFA) for workflow scheduling in cloud

computing which will optimize the number of iterations. They selected execution time, execution cost, and termination delay parameters to compare their method with GA. The proposed method optimizes the makespan, execution time, and response time.

Rajagopalan et al. [19] propounded a novel meta-heuristic algorithm of hybrid FA-GA combination for task scheduling. The proposed algorithm combines the benefits of a mathematical optimization algorithm like FA with an evolutionary algorithm like GA to form a powerful meta-heuristic search algorithm. The performance of the proposed algorithm is compared with traditional First In First Out (FIFO) and Genetic algorithms in terms of execution time.

Adil et al. [20] introduced a job scheduling mechanism based on the Discrete Firefly Algorithm (DFA) to map the grid jobs to available resources. Traditional scheduling mechanisms such as Tabu Search (TS) and hill-climbing used single-path solutions. The proposed scheduling mechanism uses population-based candidate solutions rather than single path solutions, which helps to avoid trapping at the local optimum. The authors used simulation and real workload trace to evaluate their mechanism. The performance of the proposed algorithm is compared with GA and TS in terms of makespan. Bezdan et al. [21] proposed a hybridized BA-ABC for multi-objective task scheduling problems. In the proposed method, the exploration phase of BA is enhanced by the onlooker bee search from the ABC algorithm. The performance of the proposed algorithm was compared with Chaotic Symbiotic Organisms Search (CMSOS) in terms of makespan.

Zhou et al. [22] have used GA to improve the completion time, the quality of work, and the average response time to optimize task scheduling. They used greedy algorithms and improved GA. The novel algorithm named MGGS can find an optimal solution using a fewer number of iterations. The proposed algorithm is compared with GA, Min-Min, and First Come First Service (FCFS).

Taghizadehalvandi and Kamisli [23] discussed the issue of employee shift work schedule. In this case, the goal is to minimize the total amount of employee's work and provide preferences to employees. Under this multi-objective structure, a multi-objective decision model has been created, taking into account the needs of employees. The authors developed weights/priorities of the objective functions. This study is conducted for a company operating in the service part. The number of personnel required for each department and the number of departments are the limitations of these companies. But the proposed model can be used in different types of services and work shifts.

Table 1 shows the summary of the related task scheduling algorithms in the cloud environment. From the review of the existing strategies, we can see that most of those approaches focus on minimizing the execution time or makespan of the task scheduling without regarding the QoS metrics such as throughput and energy consumption of the cloud servers. It causes load imbalance and user dissatisfaction. To increase the efficiency, the proposed algorithm finds the suitable virtual machine for each task using multiple scheduling objectives (i.e., makespan, energy consumption, throughput, tardiness, and resource utilization).

TABLE 1. Some references in task scheduling

Ref.	Year	Algorithm(s)	Compared Methods	Objective Function(s)	Weaknesses
[13]	2019	PSO	GA Greedy	Execution time	Low scalability Not considering cost and deadline
[14]	2021	MGPSOC	PSO Greedy-PSO	Makespan	Single objective
[15]	2021	GWO-PSO	GA GWO	Makespan Execution time Response time	High time complexity High overhead in scheduling
[16]	2021	ICA-PSO	PSO ICA	Makespan	Not considering energy consumption and tardiness
[17]	2021	CEFA	PSO RCT RTC RWO	Makespan Response time	Low scalability Not considering resource utilization and throughput
[18]	2021	Hybrid Cost-Effective Genetic and Firefly Algorithm (GAFFA)	GA	Finish time Execution cost Delay	High time complexity Not comparison with other meta-heuristic algorithms

[19]	2020	Hybrid Firefly-Genetic	GA FIFO	Execution time	High time complexity Not optimized QOS metrixes
[20]	2014	DFA	GA TS	Makespan	Low availability Falling in the local optimum
[21]	2021	BA-ABC	Chaotic symbiotic organisms search (CMSOS)	Makespan Cost	Load imbalance Not considering tardiness
[22]	2020	Improved GA	GA Min-Min FCFS	Makespan Response time	Load imbalance Not comparison with other meta-heuristic algorithms

3. MATHEMATICAL MODELS AND PROBLEM FORMULATION

3. 1. Problem Definition

Definition 1: (Initial preparation time for each task). Which is marked with the symbol S_0.

Definition 2: (Preparation time between tasks). Which is marked with the symbol S. In fact, S is an $N \times N$ matrix.

Definition 3: (The ratio of computational requirements to processing rate of machine ($et_{i,j}$)) [24].

$$et_{i,j} = st_i \,/\, pr_j \tag{1}$$

In Equation (1), st_i is the computational requirements of the $i-th$ task, and pr_j is the processing rate of the $j-th$ machine.

Definition 4: (The completion time of tasks in each machine (CTM)). We indicate the execution time on each machine with the symbol $ST_{i,j}$. So, we have:

$$CTM_j = \sum_{i=1}^{k} \left(et_{i,j} + ST_{i,j} \right) \tag{2}$$

In Equation (2), k *is* the number of tasks performed on the $j-th$ machine.

Definition 5: (The completion time of each task (CTT_i)). Since each task can only be performed on one machine and does not leave the machine until it is completed. CTT_i, is the time interval between $i-th$ task and $(i-1)-th$ task, with the addition of the completion time of the previous tasks.

3. 2. Objective Functions

In this section, five objective functions are introduced.

3. 2. 1. Makespan The makespan of a task scheduling depends on the execution time of each task on the selected machine instance vm_j [25].

$$makespan = \max_{1 \prec j \prec m} \left\{ CTM_j \right\} \tag{3}$$

In Equation (3), CTM_j is completion time of tasks in the $j-th$ machine.

3. 2. 2. Energy Consumption The energy consumption (EC) is the sum of energy consumed on each selected machine. It is characterized by Equation (4) [24].

$$EC_u = \varepsilon_u * \sum_{\forall i \in u} et_{i,u} \tag{4}$$

In Equation (4), EC_u is the energy consumption in the $u-th$ machine, ε_u: is the static energy consumption per time unit of the $u-th$ machine [24].

$$EC = \sum_{\forall u} EC_u \tag{5}$$

In Equation (5), EC is the total energy consumption.

3. 2. 3. Tardiness Lateness is an amount of delay in executing certain operations which is characterized by Equation (6) [26].

$$lateness_i = CTT_i - d_i \tag{6}$$

In Equation (6), CTT_i is the completion time of the $i-th$ task and d_i is the due date of the $i-th$ task. Tardiness is a measure of a delay in executing certain operations (T_i). It is characterized by Equation (7) [27].

$$T_i = \max\left(lateness_i, 0 \right) \tag{7}$$

3. 2. 4. Resource Utilization The resource utilization can be defined as the ratio between the currently used resources to the maximum resource capacity of a vm. The higher resource utilization indicates that a vm with a higher load and low utilization indicates the minimum load of a vm. Resource utilization RU_j is calculated as follows [25]:

$$RU_j = \frac{load_j}{load_j^{\max}} / 100 \tag{8}$$

In Equation (8), $load_j^{\max}$ represents the maximum permissible load of the vm_j.

$$load = std_{1 \prec j \prec m}\left\{CTM_j\right\} \quad (9)$$

In Equation (9), std is the standard deviation of completion time on the $j-th$ machine.

3. 2. 5. Throughput Throughput (TP) is the number of data outputs produced at the end task per time unit. Throughput is calculated by finding the longest or slowest task. Throughput (TP) is calculated as follows [24]:

$$TP = 1/\max\left(et_{i,j}\right) \quad (10)$$

The throughput is directly related to the longest task. The throughput of a system is low if it has at least one long task.

Throughput refers to the number of tasks completed successfully per total processing time [28]. The longest task is a bottleneck in task scheduling. When there is a big task, the finishing time of tasks increased. Therefore, when processing time is limited, fewer tasks are completed. As a result, according to TABLE 2, the throughput is reduced. In this example, it is assumed that there is only one machine. Therefore, the processing power of the machine is the same in both systems.

The general objective function is in the form of Equation (11).

$$\min F(x) = \left(w_1.Z1 + w_2.Z2 + w_3.Z3 + w_4.Z4\right)/\left(w_5.Z5\right) \quad (11)$$

$$Z1: \min\ f_1(x) = \max\left\{CTM_j\right\} \quad (12)$$

$$Z2: \min\ f_2(x) = \max\left\{EC\right\} \quad (13)$$

$$Z3: \min\ f_3(x) = \max\left\{T_i\right\} \quad (14)$$

$$Z4: \min\ f_4(x) = \max\left\{RU_j\right\} \quad (15)$$

$$Z5: \max\ f_5(x) = \max\left\{TP\right\} \quad (16)$$

The used parameters are described in Table 3.

According to Table 4, Some information about tasks and machines is provided in the form of input parameters. During the scheduling, a series of information is calculated, which are considered as calculated parameters. Finally, the optimal scheduling is the output of the system.

TABLE 2. Calculation of throughput for two assumed systems

	Size of task 1	Size of task 2	Size of task 3	Size of task 4	Throughput
System 1	12	14	11	15	0.067
System 2	11	14	19	7	0.053

TABLE 3. Symbols and definitions

N	Number of tasks
M	Number of machines
st_i	The size of the $i-th$ task
$lateness_i$	The lateness of the $i-th$ task
d_i	The due date of the $i-th$ task
T_i	The Tardiness of the $i-th$ task
pr_j	Processing rate of the *j-th* machine
ptt_i	The processing time of the $i-th$ task
CTT_i	The completion time of the $i-th$ task
S_0	Initial preparation time for each task
CTM_j	The completion time on the $j-th$ machine
S_{ab}	Preparation time between tasks ($N \times N$ matrix)
std_j	The standard deviation of completion time on the $j-th$ machine
$et_{i,j}$	The size of the $i-th$ task / the processing power of the $j-th$ machine

TABLE 4. System input, output, and calculated parameters

Inputs	Calculated	Output
N	$lateness_i$	Optimal scheduling
M	T_i	
st_i	CTM_j	
pr_j	TP	
ptt_i	std_j	
CTT_i	EC	
S_0	RU_j	
d_i		
S_{ab}		
$et_{i,j}$		

4. TECHNICAL SOLUTION

4. 1. Pathfinder Algorithm-Continuous PFA is a new swarm-based meta-heuristic algorithm that solves optimization problems with different structure [29]. This

method is inspired by collective movement of animal group and mimics the leadership hierarchy of swarms to find best food area or prey. PFA is able to converge globally optimum and avoids the local optima effectively. In this algorithm, two separated mathematical formulations are used for the position updating of the leader and other members. The mathematical formulation is used for position updating of other members that is characterized by Equation (17) [29].

$$x_i^{k+1} = x_i^k + R_1.\left(x_j^k - x_i^k\right) + R_2.\left(x_p^k - x_i^k\right) + \varepsilon \quad (17)$$

In Equation (17), k is the current iteration, x_i is the position vector of $i-th$ member, x_j is the position vector of the $j-th$ member, R_1 and R_2 are the random vectors.

$$R_1 = \alpha r_1, R_2 = \beta r_2 \quad (18)$$

In Equation (18), α is the coefficient for interaction which defines the magnitude of movement of any member together with its neighbor and β is the coefficient of attraction which sets the random distance for keeping the herd roughly with leader. In this study, α and β are randomly selected in the range of [1, 2].

Also, r_1 and r_2 provide a random movement and are uniformly generated in the range of [0,1]. ε is for vibration that is generated in each iteration using Equation (19).

$$\varepsilon = \left(1 - \frac{k}{k_{\max}}\right).u_1.D_{ij}, D_{ij} = \left\|x_i - x_j\right\| \quad (19)$$

In Equation (19), u_1 is random vectors range in $[-1,1]$, D_{ij} is the distance between two members and $k_{\max}$ is the maximum number of iterations. To look for prey, the mathematical formulation is used for position updating of the pathfinder characterized by Equation (20).

$$x_p^{k+1} = x_p^k + 2r_3.\left(x_p^k - x_p^{k-1}\right) + A \quad (20)$$

In Equation (20), r_3 is a random vector that is uniformly generated in the range of [0,1], A is generated in each iteration using Equation (21).

$$A = u_2.e^{\frac{-2k}{k_{\max}}} \quad (21)$$

where, u_2 is random vectors range in $[-1,1]$.

4. 2. Discretization Technique The task scheduling problem is considered as a discrete optimization problem and PFA is proposed to deal with a continuous optimization problem. Therefore, in this paper, this issue can handle by initially generating solutions using Equation (22) [30].

TABLE 5. Pseudo code of PFA [29].

```
Load PFA parameter
Initialize the population
Calculate the fitness of initial population
Find the pathfinder
While k < k_max
    α and β = random number in [1,2]
    Update the position of pathfinder using Equation (20) and and check the bound
    If new pathfinder is better than old
        Update pathfinder
    end
    for i=2 to maximum number of populations
        Update the position of members using Equation (17) and and check the bound
    end
    calculate new fitness of members
    find the best fitness
    If best fitness< fitness of pathfinder
        Pathfinder=best members
        Fitness= best fitness
    end
    for i=2 to maximum number of populations
        If new fitness of member(i)<fitness of member (i)
            Update members
        end
    end
    generate new A and ε
end
```

$$X_{ij} = floor\left(Lb_{ij} + \alpha.\left(Ub_{ij} - Lb_{ij}\right)\right), \alpha \in [0,1], j = 1,2,...,n \quad (22)$$

In Equation (22), the lower boundary Lb is set to 1, while the upper boundary Ub is set to M. Figure 3 shows the main loop of the proposed algorithms. Discretization in the main loop does not add time complexity.

4. 2. Time Complexity Similar to population-based algorithms such as PSO and GOA, the proposed algorithm has an iteration loop in which some operations are performed according to the dimensions of the problem and the objective functions. Consequently, the computational complexity of the proposed method is $O\left(t\left(n * PS + \left(F\right) * PS\right)\right)$, where t is the number of iterations, n is the number of dimensions, PS is the population size of swarm, and F is the cost of objective [29]. This amount of time consumed to run the algorithm is reasonable and common.

4. 3. Implementation and Parameter Setting First, we determine the number of tasks, the number of machines, the size and time of execution of tasks, as well as the processing power of machines. Then, to solve the task scheduling problem using meta-heuristic algorithms, we need to create an initial population. To create the initial population, we must determine the number of tasks

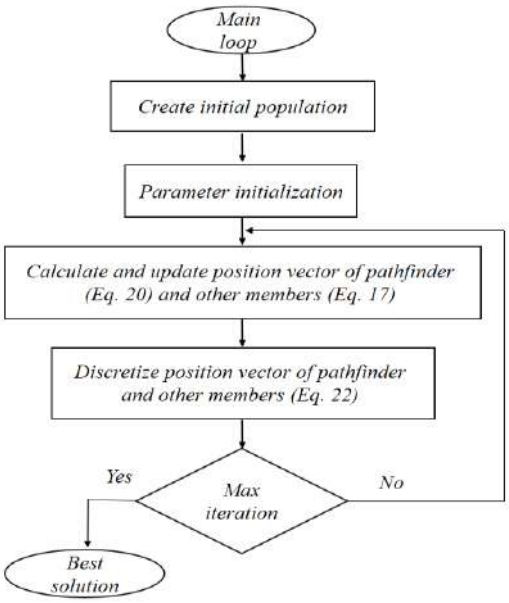


Figure 3. Flowchart of the proposed algorithm

and the number of machines. The *x* vector is a sample of the population shown in Equation (23):

$$x = \{4,3,1,3,2\} \quad (23)$$

Correspond to this representation, the first task will be accepted on the fourth machine. The second and fourth tasks are accepted on the third machine, while the third task will be accepted on the first machine, the fifth task will be accepted on the second machine. In this paper, we use FA, BA, PSO, and GOA algorithms for comparison. The adjustable parameters of these methods are introduced in Tables 6 to 10. MATLAB software has been used for simulation. First, we create several models, the number of machines is 20 and the number of tasks is 40, 60, and 120. Each of the algorithms is executed 10 times, the number of particles is 50, the max number of iteration is 100. The average results and the worst/best result are given in Tables 11 to 13.

5. PERFORMANCE EVALUATION

This section is dedicated to simulate algorithms and present results. In all tables, bold numbers mean the best

TABLE 6. Adjustable parameters of BA

Loudness	0.9
Pulse rate	0.1
Frequency minimum	0
Frequency maximum	5

TABLE 7. Adjustable parameters of PSO

$P1$	2.05
$P2$	2.05
P	$P1+P2$
C	$2/\left(P-2+\sqrt{P^2-4P}\right)$
D	0.3
Inertia weight	C
Inertia Weight Damping Ratio	0.1
Social learning factor $c1$	$C\times P1$
Personal learning factor $c2$	$C\times P2$
Velocity of the particles	Range of vm_s

TABLE 8. Adjustable parameters of FA

Light Absorption Coefficient	2
Attraction Coefficient Base Value	2
Mutation Coefficient	0.2

TABLE 9. Adjustable parameters of GOA

cMin	0.00001
cMax	1

TABLE 10. Adjustable parameters of DPFA

Alpha	Random number in [0,1]
Beta	Random number in [0,1]

results of the algorithms. W, B, and M character indicates worst, best, and average case, respectively.

In Table 11, scheduling for 40 tasks is considered. The results shows that DPFA for *Z1*, *Z2*, and *Z5* objective functions with an average of 56.64, 706.13, and 0.018, respectively, is the best among the other. Moreover, DPFA has the best result for the total objective function (i.e., *Z* that is obtained by Equation (11)). The performance gap between DPFA and FA, BA, PSO, and GOA is 3.66%, 77.50%, 6.40%, and 3.58%, respectively.

Table 12 shows the results of the algorithms for scheduling 60 tasks. DPFA for *Z3*, *Z4*, and *Z5* objective functions with an average of 717.00, 5.09, and 0.010, respectively, is the best among the other algorithms. With these interpretations, DPFA has better results than others. GOA is better in terms of run-time, but because DPFA achieves the optimal answer in fewer iterations, it can be concluded that it is better.

Table 13 shows the results of the algorithms for scheduling 120 tasks. The DPFA for *Z1* objective

TABLE 11. Comparison of algorithms for scheduling 40 tasks on 20 machines based on five objective functions

		FA	BA	PSO	GOA	DPFA
	W	**66.19**	190.47	74.30	91.65	84.33
Z1	**B**	47.99	49.32	47.55	47.15	**43.42**
	M	57.21	91.52	65.49	68.41	**56.64**
	W	836.91	832.70	851.75	808.20	**804.90**
Z2	**B**	689.99	620.64	**605.65**	190.86	617.62
	M	749.70	715.29	723.71	519.57	**706.13**
	W	445.90	376.60	377.50	466.85	**302.15**
Z3	**B**	279.15	208.00	240.00	200.90	**150.10**
	M	329.82	279.37	317.13	**219.11**	248.97
	W	6.75	6.68	6.96	**6.01**	6.49
Z4	**B**	5.82	**3.95**	5.83	4.45	5.50
	M	6.33	5.65	6.39	**5.48**	6.11
	W	**0.018**	0.005	0.013	0.010	0.011
Z5	**B**	0.025	0.020	0.023	0.021	**0.028**
	M	0.014	0.012	0.015	0.015	**0.018**
Z		59404.9	101721.6	60977.5	59359.0	**57309.0**
Time		3.8	3.5	3.2	**3.0**	3.5

TABLE 12. Comparison of algorithms for scheduling 60 tasks on 20 machines based on five objective functions

		FA	BA	PSO	GOA	DPFA
	W	508.95	191.05	**111.99**	134.92	116.63
Z 1	**B**	68.80	95.28	71.65	70.85	**68.60**
	M	157.26	135.57	**89.48**	104.69	98.95
	W	1661.71	1759.19	**1551.18**	1835.59	1703.15
Z 2	**B**	1325.53	1343.39	**1220.76**	1063.87	1363.11
	M	1450.70	1500.17	**1373.54**	1577.08	1453.11
	W	1147.30	944.60	1024.75	1059.10	**955.70**
Z 3	**B**	707.95	613.10	657.65	**489.10**	587.55
	M	855.13	734.79	828.76	771.56	**717.00**
	W	6.44	6.12	6.53	**6.12**	6.08
Z 4	**B**	**2.81**	4.29	5.59	4.56	5.18
	M	5.33	5.20	5.98	5.24	**5.09**
	W	0.001	0.005	0.008	0.007	**0.008**
Z 5	**B**	0.010	0.010	0.013	0.014	**0.014**
	M	0.009	0.007	0.010	0.010	**0.010**
Z		415426.2	323743.9	216106.9	207979.9	**205710.3**
Time		4.8	4.1	4.8	**3.9**	5.8

function with an average of 422.17, for *Z2* objective function with an average of 4903.01, for *Z3* objective function with an average of 4140.54, for *Z5* objective function with an average of 0.002 has great efficiency. As a result, the DPFA has a 0.51%, 31.10%, 0.27%, and 5.29% improvement in results compared to the FA, BA, PSO, and GOA, respectively. The high run-time of DPFA is compensated by reaching the solution in a low number of iteration.

TABLE 13: Comparison of algorithms for scheduling 120 tasks on 20 machines based on five objective functions

		FA	BA	PSO	GOA	DPFA
	W	709.05	774.90	**513.25**	547.60	565.95
Z 1	**B**	**264**.62	377.10	368.65	340.45	**313.48**
	M	416.21	572.77	443.56	424.52	**422.17**
	W	6650.81	6249.32	**6410.62**	6410.63	6166.46
Z 2	**B**	4684.75	4842.42	4813.98	4517.74	**4353.34**
	M	5660.80	5671.36	5683.41	4938.35	**4903.01**
	W	4946.20	4973.20	4502.05	4917.95	**4295.50**
Z 3	**B**	3994.00	4048.85	3996.00	**3480.00**	4018.40
	M	4398.53	4329.38	4248.35	4241.53	**4140.54**
	W	4.78	4.26	4.24	4.76	**4.03**
Z 4	**B**	3.18	**3.08**	3.74	3.48	3.57
	M	4.10	**3.63**	3.94	3.91	4.11
	W	0.001	0.001	0.001	0.001	**0.001**
Z 5	**B**	0.003	0.002	0.002	0.003	**0.003**
	M	0.002	0.001	0.002	0.002	**0.002**
Z		4634241.3	6044950.7	4623503.0	4854967.3	**4610850.0**
Time		**2.4**	5.3	6.6	6.1	7.9

TABLE 14. Some weakness of optimization algorithms

Algorithm	weakness
FA	-Easy to fall into local optimum [31] -The solution accuracy is lower -Sticking to local minima [32]
BA	-The low precision of optimization [33] -Easy to fall into local optimum -A poor iterative ability -Its global searchability is weak -It is difficult to process high-dimensional data
PSO	-Easy to fall into local trapping [34] -Premature convergence
GOA	-Easy to fall into local optimum [35] -Slow convergence speed

In Table 14, some weaknesses of the algorithms used in this paper are listed.

Figure 4 compares the algorithms in terms of makespan. The average makespan value is gradually increased with increasing the size of tasks. In a small number of tasks, the algorithms performed almost similarly. As the number of tasks is increased, FA, GOA, and DPFA can complete tasks in less time. Among the algorithms, GOA and DPFA have the best performance. DPFA, in 40, 60, and 120 tasks has a suitable result with minimal fluctuations. Having two separate equations to update the position of the herd members makes the DPFA algorithm works properly. Compared with other optimization algorithms, BA has a poor iterative ability.

Figure 5 compares the algorithms in terms of energy consumption. The average energy consumption value is gradually increased with increasing the number of tasks. The best performance for 40 tasks is related to GOA and the best performance for 60 tasks is related to PSO. When we have 40 or 60 tasks, the other algorithms performed almost similarly. As the number of tasks increased,

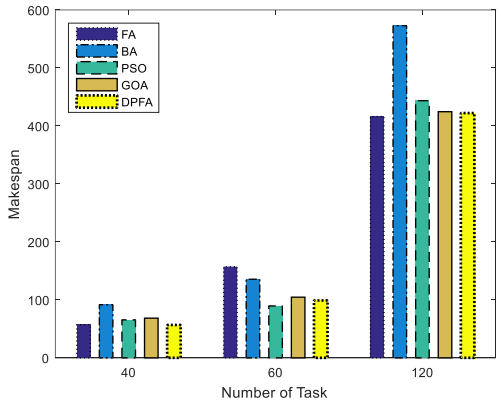


Figure 4. Algorithms' comparison based on makespan

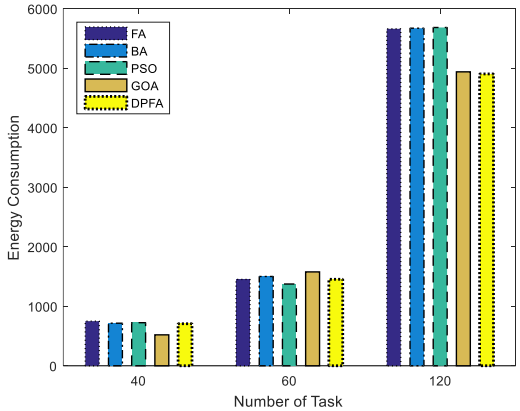


Figure 5. Algorithms' comparison based on energy consumption

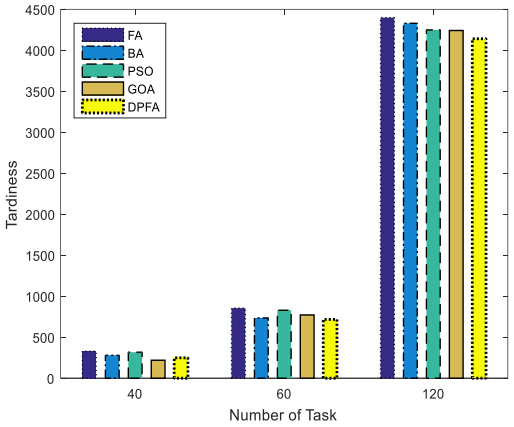


Figure 6. Algorithms' comparison based on tardiness

DPFA can complete tasks with less energy consumption and show the best performance. The high convergence speed of the DPFA causes it to converge in fewer iterations and reduces energy consumption.

Tardiness is the other parameter that is added to the objective function in this paper. Naturally, a small amount of this parameter makes the response time shorter. Among the algorithms, DPFA with the lowest latency is able to have better results than the rest of the algorithms (Figure 6). One of the important factors of DPFA is its speed and simplicity, which has appeared in this result.

Figure 7 compares the algorithms in terms of resource utilization. The lower value of this parameter is better and low utilization indicates the minimum load of a VM. The best performance for 40 tasks is related to GOA, the best performance for 60 tasks is related to DPFA, and the best performance for 120 tasks is related to BA. But on average and due to Table 13, he DPFA is better than the other algorithms.

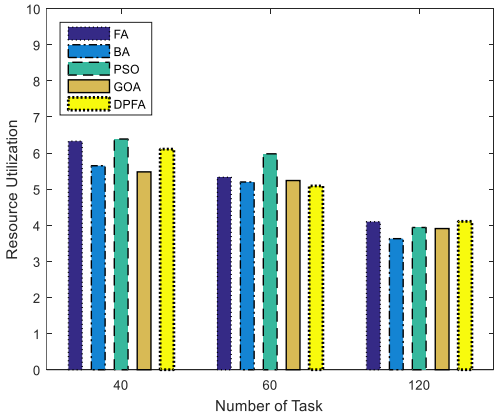


Figure 7. Algorithms' comparison based on resource utilization

Figure 8 compares the algorithms in terms of throughput. The higher value of this parameter is better. The best performance for 40, 60, and 120 tasks is related to DPFA. The DPFA has a high exploration capability in a low number of iterations, so it assigns suitable machines to the tasks and throughput increases. The worst performance for 40, 60, and 120 tasks is related to BA, because it has some shortcomings, such as its global searchability is weak and it is difficult to process high-dimensional data.

The number of iterations is an effective parameter for all algorithms. Figure 9 shows the proper performance of the FA algorithm. In 50 iterations, this algorithm has a completion time of 585426.2. It reduces this time to 233617.0 in 450 iterations that means a 60% improvement. In this respect, it has the largest reduction compared to other algorithms. DPFA does not change with an increasing number of iterations. If the number of iterations is increased, finding of new promising solutions will be difficult because of fluctuation rate A

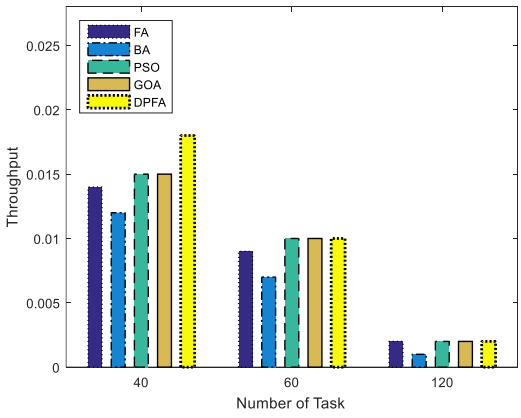


Figure 8. Algorithms' comparison based on throughput

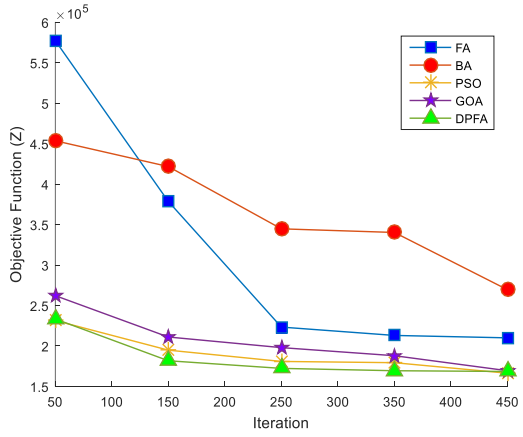


Figure 9. Algorithms' comparison based on iteration and objective function (Z) (60 tasks and 20 machines- each algorithm is run twice and averaged)

and vibration vector ε converging to 0. In fact, the convergence speed of this algorithm has appeared with very good results in a small number of iterations, and increasing the number of iterations does not change the result.

6. CONCLUSIONS

The task scheduling problem is one of the most critical issues in cloud computing because cloud performance depends mainly on it. The main aim of a novel scheduling technique is to evaluate the optimum set of resources available to execute an incoming task such that a scheduling algorithm (scheduler) can then be applied to optimize such diverse QoS parameters as cost, makespan, scalability, reliability, resource utilization, energy consumption, etc.

In this paper, we modeled the objective function based on five parameters (i.e., makespan, energy consumption, tardiness, resource utilization, and throughput). We used a meta-heuristic algorithm, named PFA and proposed a new task scheduling algorithm based on Discrete PFA (DPFA). DPFA has some features (i.e., has two separate mathematical formulations for position updating of head and other members- can explore promising solutions- provide the abrupt changes in the initial iterations- exploit the best one throughout iterations). Due to these reasons, when the number of tasks increases, it has the best performance for makespan, energy consumption, tardiness, and throughput objective functions. On average, among the algorithms, DPFA has the desirable performance and can optimize the total objective function. We suggest using chaos theory to create diversity in the population, adjusting parameters of PFA with the help of tools (i.e., fuzzy system). It can also be implemented to other problems in different areas.

7. REFERENCES

1. Helo. P, Hao. Y, Toshev. R, and Boldosova. V, "Cloud manufacturing ecosystem analysis and design", ***Robotics and Computer Integrated Manufacturing***, (2021), doi:10.1016/j.rcim.2020.102050.
2. Mohammad Hasani zade. B, Mansouri. N, and Javidi. MM, "Multi-objective scheduling technique based on hybrid hitchcock bird algorithm and fuzzy signature in cloud computing", ***Engineering Applications of Artificial Intelligence***, Vol. 104, (2021), doi: 10.1016/j.engappai.2021.104372.
3. Mansouri. N, "An effective weighted data replication strategy for data Grid", ***Australian Journal of Basic and Applied Sciences***, Vol. 6, No. 10, (2012), 336-346.
4. Wang. B, Wang. Ch , Huang. W, Song. Y, and Qin. X, "Security-aware task scheduling with deadline constraints on heterogeneous hybrid clouds", ***Journal of Parallel and Distributed Computing***, Vol. 153, (2021), 15-28, doi: 10.1016/j.jpdc.2021.03.003.
5. Yang. XS, "Firefly algorithms for multimodal optimization", ***Stochastic Algorithms: Foundations and Applications***, (2009), 169-178, doi: 10.1007/978-3-642-04944-6_14.

6. Kennedy. J, and Eberhart. RC, "Particle swarm optimization", ***IEEE International Conference on Neural Networks (Perth, Australia)***, Piscataway, Vol. 5, (1995), 1942-1948, doi: 10.1007/s11721-007-0002-0.
7. Yang. XS, "A new metaheuristic bat-inspired algorithm", ***Studies in Computational Intelligence***, Vol. 284, (2010), 65-74. https://doi.org/10.1007/978-3-642-12538-6_6.
8. Saremi. S, Mirjalili. S, and Lewis. A, "Grasshopper optimisation algorithm: theory and application", ***Advances in Engineering Software***, Vol. 105, (2017), 30-47, doi: 10.1016/j.advengsoft.2017.01.004.
1. 9. Nicolas. A, Rafael. R, David. F, and Raul. P, "A discrete firefly algorithm for solving the flexible job-shop scheduling problem in a make-to-order manufacturing system", ***Central European Journal of Operations Research,*** (2020), doi: 10.1007/s10100-020-00701-w.
9. Kumar. M, and Sharma. SC, "PSO-based novel resource scheduling technique to improve QoS parameters in cloud computing", ***Neural Computing and Applications,*** Vol. 32, (2020), 12103-12126, doi: 10.1007%2Fs00521-019-04266-x.
10. Mansouri. N, Ghafari. R, and Mohammad Hasani Zade B, "Cloud computing simulators: A comprehensive review", ***Simulation Modelling Practice and Theory***, Vol. 104, (2020), doi: 10.1016/j.simpat.2020.102144.
11. Shareh. MB, Bargh. SH, Hosseinabadi. AAR, and Slowik. A, "An improved bat optimization algorithm to solve the tasks scheduling problem in open shop", ***Neural Computing and Applications***, Vol. 33, (2021), 1559-1573, doi: 10.1007/s00521-020-05055-7.
12. Agarwal. M, and Srivastava. GMS, "A PSO algorithm-based task scheduling in cloud computing**",** ***International Journal of Applied Metaheuristic Computing (IJAMC)***, Vol. 10, (2019), doi: 10.4018/IJAMC.2019100101.
13. Raju. YHP, and Devarakonda. N, "Greedy-based PSO with clustering technique for cloud task scheduling", Proceedings of International Conference on Computational Intelligence and Data Engineering, (2021), 133-141, doi: 10.1007/978-981-15-8767-2_12.
14. Kumar. SMS, Krishnamoorthy. P, Soubraylu. S, Venugopal. And JK, Marimuthu. K, "An efficient task scheduling using GWO-PSO algorithm in a cloud computing environment", Proceedings of International Conference on Intelligent Computing, Information and Control Systems, (2021), 751-761, doi: 10.1007/978-981-15-8443-5_64.
15. Tabrizchi. H, Rafsanjani. MK, Balas VE., "Multi-task scheduling algorithm based on self-adaptive hybrid ICA–PSO algorithm in cloud environment", ***Soft Computing Applications,*** (2021), 422-431, doi: 10.1007/978-3-030-52190-5_30.
16. Chakravarthi. KK, Shyamala. L, Vaidehi. V, "Cost-effective workflow scheduling approach on cloud under deadline constraint using firefly algorithm", ***Applied Intelligence***, Vol. 51, (2021), 1629-1644, doi: 10.1007/s10489-020-01875-1.
17. Kaur. I, and Mann. PS, "A hybrid cost-effective genetic and firefly algorithm for workflow scheduling in cloud", International Conference on Innovative Computing and Communications, (2021), doi: 10.1007/978-981-15-5148-2_4.
18. Rajagopalan A, Modale. DR, and Senthilkumar. R, "Optimal scheduling of tasks in cloud computing using hybrid firefly-genetic algorithm", ***Advances in Decision Sciences, Image Processing, Security and Computer Vision***, (2020), 678-687, doi: 10.1007/978-3-030-24318-0_77.
19. Yousif. A, Nor. SM, Abdullah. AH, and Bashir. MB, "A discrete firefly algorithm for scheduling jobs on computational grid", ***Studies in Computational Intelligence***, (2014), 271-290, doi: 10.1007/978-3-319-02141-6_13.
20. Bezdan. T , Zivkovic. M , Tuba. E , Strumberger. I , Bacanin. N, and Tuba. M, "Multi-objective task scheduling in cloud computing environment by hybridized bat algorithm", ***Advances in Intelligent Systems and Computing***, (2021), 718-725, doi: 10.1007/978-3-030-51156-2_83.
21. Zhou Zh, Li. F, Zhu. H, and Xie. H, Abawajy. JH, Chowdhury M U., "An improved genetic algorithm using greedy strategy toward task scheduling optimization in cloud environments", ***Neural Computing and Applications***, Vol. 32, (2020), 1531-1541, doi: 10.1007/s00521-019-04119-7.
22. Taghizadehalvandi. M, and Kamisli. OZ, "Multi-objective Solution Approaches for Employee Shift Scheduling Problems in Service Sectors", ***International Journal of Engineering, Trnsactions C: Aspects,*** Vol. 32, (2019), 1312-1319, doi: 10.5829/ije.2019.32.09c.12.
23. Gu. Y, and Budati C, "Energy-aware workflow scheduling and optimization in clouds using bat algorithm", ***Future Generation Computer Systems***, Vol. 113, (2020), 106-112, doi: 10.1016/j.future.2020.06.031.
24. Adhikari. M, Amgoth. T, and Srirama. SN, "Multi-objective scheduling strategy for scientific workflows in cloud environment a firefly-based approach," ***Applied Soft Computing Journal***, Vol. 93, (2020), doi: 10.1016/j.asoc.2020.106411.
25. Su S, and Yu. H, "Minimizing tardiness in data aggregation scheduling with due date consideration for single-hop wireless sensor networks", ***Wireless Networks***, Vol. 21, (2015), 1259-1273, doi: 10.1007/s11276-014-0853-4.
26. Akyol. DE, and Bayhan. GM, "Multi-machine earliness and tardiness scheduling problem: an interconnected neural network approach", ***The International Journal of Advanced Manufacturing Technology*** , Vol. 37, (2008), 576-588, doi: 10.1007/s00170-007-0993-0.
27. Kumar. M, and Sharma. SC, "PSO-COGENT: Cost and energy efficient scheduling in cloud environment with deadline constraint", ***Sustainable Computing: Informatics and Systems***, Vol. 19, (2018), 147-164, doi: 10.1016/j.suscom.2018.06.002.
28. Yapici. H, and Cetinkaya. N, "A new meta-heuristic optimizer: pathfinder algorithm", ***Applied Soft Computing Journal***, Vol. 78, (2019), 545-568, doi: 10.1016/j.asoc.2019.03.012.
29. Elaziz. MA, and Attiya. I, "An improved henry gas solubility optimization algorithm for task scheduling in cloud computing", ***Artificial Intelligence Review***, (2020), doi: 10.1007/s10462-020-09933-3.
30. Panda. MR, Panda. S, Priyadarshini. R, and Das. P, "Mobile robot path-planning using oppositional-based improved firefly algorithm under cluttered environment", ***Advances in Intelligent Computing and Communication***, (2020), doi: 10.1007/978-981-15-2774-6_18.
31. Yu. Y, Ren. X, Du. F, and Shi. J, "Application of improved PSO algorithm in hydraulic pressing system identification", ***Journal of Iron and Steel Research, International***, Vol. 19, (2012), 29-35, doi: 10.1016/S1006-706X(13)60005-9.
32. Li. X, Zhou. L, Deng. X, Wang. B, Qiu. C, Lu. M, and Li. C, "A resource allocation scheme based on improved bat algorithm for D2D communication system", International Conference in Communications, Signal Processing, and Systems CSPS 2018: Communications, Signal Processing, and Systems, Vol. 515, (2019), 852-861, doi: 10.1007/978-981-13-6264-4_100.
33. Song. B, Wang. Z, and Zou. L, "An improved PSO algorithm for smooth path planning of mobile robots using continuous high-degree bezier curve", ***Applied Soft Computing Journal***, Vol. 100, (2021), doi: 10.1016/j.asoc.2020.106960.
34. Luo. J, Chen. H, Zhang. Q, Xu. Y, Huang. H, and Zhao. X, "An improved grasshopper optimization algorithm with application to financial stress prediction", ***Applied Mathematical Modelling***, Vol. 64, (2018), 654-668, doi: 10.1016/j.apm.2018.07.044.

Persian Abstract

چکیده

زمان‌بندی کار یکی از موضوعات اساسی می‌باشد که توجه بسیاری از محققان را به منظور افزایش عملکرد ابر و رضایت مصرف‌کننده جلب کرده است. زمان‌بندی کار یک مسئله‌ی NP سخت است که به دلیل وجود چندین هدف متناقض، برای کاربران و ارائه دهندگان خدمات، چالش برانگیز است. بنابراین، الگوریتم‌های فراابتکاری برای حل مسئله‌ی زمان‌بندی کار در زمان قابل قبول، گزینه‌ی مناسب‌تری هستند. اگرچه الگوریتم‌های زمان‌بندی کار بسیاری ارائه شده‌اند، اما رویکردهای موجود عمدتا روی به حداقل رساندن زمان اتمام کارها یا میزان مصرف انرژی، متمرکز شده و سایر پارامترهای مهم را نادیده گرفته‌اند. در این مقاله، ما الگوریتم جدیدی را برای زمان‌بندی کار مبتنی بر الگوریتم بهینه‌سازی مسیریاب گسسته (DPFA)، پیشنهاد می‌دهیم که از حرکت جمعی گروه حیوانات الهام گرفته و از سلسله مراتب دسته‌ها برای یافتن شکار تقلید می‌کند. زمان‌بند پیشنهادی پنج هدف (به عنوان مثال، زمان اتمام کارها، مصرف انرژی، توان عملیاتی، تأخیر و استفاده از منابع) را به عنوان توابع هدف در نظر گرفته است. در نهایت، الگوریتم‌های مختلفی مانند الگوریتم کرم شب‌تاب (FA)، الگوریتم بهینه‌سازی ازدحام ذرات(PSO)، الگوریتم بهینه‌سازی خفاش(BA) و الگوریتم بهینه‌سازی ملخ(GOA)، برای مقایسه استفاده شده است. نتایج تجربی نشان می‌دهد که الگوریتم زمان‌بندی پیشنهادی برای تابع هدف زمان اتمام کارها، تا ۹/۱۶ درصد، ۳۸/۴۴ درصد، ۳/۵۹ درصد و ۳/۴۴ درصد به ترتیب در مقایسه با الگوریتم کرم شب‌تاب، الگوریتم بهینه‌سازی ازدحام ذرات، الگوریتم بهینه‌سازی خفاش و الگوریتم بهینه‌سازی ملخ بهبود داشته است. علاوه بر این، نتایج، بهبود چشمگیر از نظر استفاده از منابع، توان عملیاتی، تاخیر و مصرف انرژی را نشان می دهد.

IJE TRANSACTIONS C: Aspects, Vol. 34, No. 09, (September 2021) 2137-2147

International Journal of Engineering

Journal Homepage

Numerical/Experimental Studies on Performance at Low Engine Speeds: A Case study Downsized Iranian National Engine (EF7)

M. M. Namar[a], O. Jahanian*[a], R. Shafaghat[a], K. Nikzadfar[a,b]

[a] *Faculty of Mechanical Engineering, Babol Noshirvani University of Technology, Babol, Iran*
[b] *School of Mechanical, Aerospace and Automotive Engineering, Coventry University, Coventry, UK*

PAPER INFO

Paper history:
Received 14 April 2021
Received in revised form 19 July 2021
Accepted 30 July 2021

Keywords:
Engine Downsizing
Performance
3-Cylinder
Turbocharger
Continuous Variable Valve Timing

ABSTRACT

Engine downsizing is a trusted method to reduce fuel consumption and pollution emitted from internal combustion engines. In this method, engine displacement volume is reduced while maintaining the same power/torque characteristics. However, there still exist several limitations to utilize this technology. In this paper, the naturally aspirated type of Iranian national engine (EF7-NA) for a possible downsized version is investigated. A one-dimensional engine model equipped with a zero-dimensional two-zone combustion sub-model was developed and validated via experimental results for both natural aspirated and turbocharged engine types. Then experimental and numerical studies were carried out for the primary concept, deactivation of one cylinder besides using a turbocharger. To overcome the concept shortages, especially in lower ranges of engine speed, numerical studies were extended. Deployment of several turbochargers with different performance maps and different valve timing via a dual continuous variable valve timing (CVVT) system were investigated. The results showed that there is a feasible method for EF7 engine downsizing via a 3-cylinder type equipped with a modified turbocharger and valve timing. The maximum difference between base-engine and downsized version torque is about 7% in low engine speeds.

***doi**: 10.5829/ije.2021.34.09c.11*

NOMENCLATURE

1D	One Dymensional	NA	Naturally Aspirated
BMEP	Brake Mean Effective Pressure	TC	Turbo-Charged
BSFC	Brake Specific Fuel Consumption	VVT	Variable Valve Timing
CR	Compression Ratio	**English Symbols**	
CVVT	Continuous Variable Valve Timing	A	Area, m^2
DI	Direct Injection	B	Bore, m
EGR	Exhaust Gas Recirculated	E	Internal Energy, kJ
EMS	Engine Management System	$\dot{m}$	Mass Flow Rate, kg/s
EVO	Exhuast Valve Opening	N	Engine Speed, rpm
GDI	Gasoline Direct Injection	T	Temperature, K
ICE	Internal Combustion Engines	V	Volume, m^3
IEA	International Energy Agency	**Greek Symbols**	
IVC	Inlet Valve Closing	θ	Crank Angle, deg
LHV	Low Heating Value	ρ	Density, kg/ m^3
LTC	Low Temparature Combustion	$\emptyset$	Equivalence Ratio

1. INTRODUCTION

Engine Downsizing is well-known as a promising approach achieving the long-term goal of the International Energy Agency (IEA) [1] focusing on a 50% reduction in mean global emission by 2030. Besides reducing the conventional emissions of Internal Combustion Engines (ICE), engine downsizing leads to

*Corresponding Author Institutional Email: jahanian@nit.ac.ir (O. Jahanian)

Please cite this article as: M. M. Namar, O. Jahanian, R. Shafaghat, K. Nikzadfar, Numerical/Experimental Study on Downsized Iranian National Engine (EF7) Performance at Low Engine Speeds, International Journal of Engineering, Transactions C: Aspects, Vol. 34, No. 09, (2021) 2137-2147

less carbon dioxide emission in the automotive industry, which is one of the main concerns for researchers who worked on global warming [2-4]. In this approach, the displacement volume of the engine is declined while the performance kept fixed. It is an effective solution to performance enhancement [5, 6], emission reduction [7], and frictional losses decrease [8]. Although extended researches from other aspects such as employing additives [9, 10] and Low Temperature Combustion (LTC) [11-13] were carried out to achieve IEA 2030 goals. In recent years, engine downsizing approach was succeeded in attaining outstanding emission reduction (32%) at European markets [14].

In general, downsized engines are designed based on reduction in the bore, connecting rod length, and/or the number of cylinders. It is obvious that to retaliate the performance, utilizing the boost technologies such as supercharger, turbocharger, Direct Injection (DI), and Variable Valve Timing (VVT) is essential. However, running these technologies may lead some challenges namely; knock, super-knock, pre-ignition, and also electrification [15-17].

The range of employed techniques for engine downsizing can be considered from redesigning the engine components to general new engine designing or even the Engine Management System (EMS). Although, in the recent years, the researches focus is shifted to coping the operational challenges and performance improvement rather than redesigning [18, 19]; it seems that there is still a great potential for ultra-downsizing yet [20].

A 50% downsized 3-cylinder engine optimal designing has been reported by Hancock et al. [21], focusing on design structure and employed technologies, and 30% fuel economy plus CO_2 reduction are reported as the results of their optimal design. The concept of using pneumatic hybridization instead of electric ones for ultra-downsizing was reported more cost-efficient by Dönitz et al. [22]. The opportunity of a spark ignition engine 40% downsizing employing high octane bio-fuels and cooled Exhaust Gas Recirculated (EGR) was investigated by Splitter and Szybist [23]. Furthermore, charge cooling with a tracer-based two-line Planar Laser Induced Fluorescence (PLIF) technique in an optical Gasoline Direct Injection (GDI) engine was introduced as an idea to increase volumetric efficiency and Compression Ratio (CR) for downsized engine by Anbari et al. [24]. In addition, Turner et al. [25] have achieved 35% CO_2 reduction by designing a 60% downsized engine from a 5L, 8-cylinder V-type Jaguar Land Rover engine. More efficient turbulent flow at intake port in part load operation was achieved using a new design of intake system by Millo et al. [26]; while it was not realized at full load. Cooperating of this new design via advancing of Inlet Valve Closing (IVC) and employing turbocharger was introduced as an effective way to improve SI engines performance. In 2015, Severi et al. [27] asserted that 20% displacement volume reduction besides providing right-size engine maximum power of studied GDI engine is achievable via 11% piston bore reduction and using both engine boosting and spark advancing, in a numerical investigation. The concept of designing a boosted uni-flow scavenged direct injection gasoline engine to achieve more than 50% downsizing is demonstrated by Ma and Zhao [28] for a two stroke engine. Furthermore, friction loss investigation due to employing micro-geometry piston bearing [29] and oil pan design for modern downsized engine [30] were studied.

TABLE 1. The domain of recent-employed strategies on engine downsizing

Strategy		Result	Reference
Engine Design	3 cylinder, Optimized design structure	30% fuel economy, CO_2 reduction	Hancock et al. [21]
	Engine redesign	35% CO_2 reduction, 60% downsizing	Turner et al. [25]
	Bore reduction	20% downsizing	Severi et al. [27]
Engine Components Design	Micro-geometry piston bearing	Frictional losses reduction, Torque enhancement	Wróblewski et al. [29]
	Oil pan design	Noise reduction	Liu at al. [30]
	Intake system	Part load operation improvement	Millo et al. [26]
Engine Management	EMS, Supercharger	35% CO_2 reduction, 60% downsizing	Turner et al. [25]
	IVC advancing, Turbocharger	Part load operation improvement	Millo et al. [26]
Other	Bio-fuels, Cooled EGR	40% downsizing	Splitter and Szybist [23]
	Boosted uni-flow scavenged GDI	>50% downsizing	Ma and Zhao [28]
	Charge cooling with PLIF	Volumetric efficiency enhancement, CR increase	Anbari et al. [24]
	Pneumatic hybridization	Cost-efficient ultra-downsizing	Dönitz et al. [22]

This brief review showed that the employed methods on engine downsizing include a wide range of strategies. Recent-employed strategies are summarized in Table 1.

Investigation in the literature clearly showed the important role of engine downsizing on emission reduction. In the other hand, considering the trend line of the Iran-domestic industries besides the daily-increasing demand of releasing clean engines, may strongly portrait the need of downsized engines in Iran markets. In consequence, in this research, the Iranian national engine (EF7) was investigated for a feasibility study of downsizing. The Naturally Aspirated (NA) 4 cylinder engine was considered as the base engine, and a 25% downsized version was introduced using the cylinder reduction strategy (3 cylinder engine) besides utilizing turbocharger and VVT technologies. To achieve this aim, a one-dimensional model which is able to investigate the base engine performance as well as applying turbocharger, cylinder reduction, and VVT is employed. The model is equipped with a zero-dimensional combustion sub-model. Indeed, to evaluate the validity of numeric results, and experimental estimation of engine performance in 3-cylinder condition, the fuel cut-off strategy is applied to the one cylinder of 4 cylinder EF7 Turbo-Charged (TC) engine. Furthermore, the model results in fuel cut-offed mode were compared via experimental results for double validity check. The methodology of this research is shown in Figure 1.

2. MODEL DESCRIPTOPN

The investigation of mass and energy equations for each component, is mainly used in engine one dimensional (1D) simulation [31]. The engine performance is finally estimated by the combination of 1D simulation results and a zero-dimensional two-zone thermodynamic simulation of the combustion process. So, the model has illustrated in two subsections namely; component and combustion simulations. The schematic of operating model is shown in Figure 2.

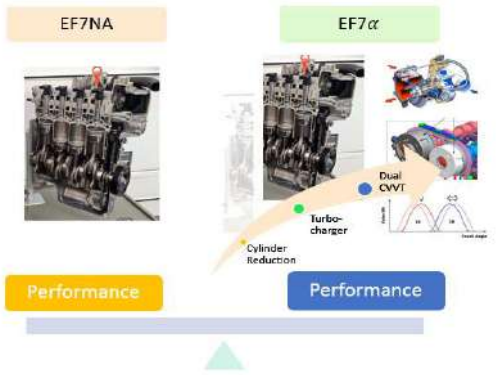


Figure 1. The methodology of this research

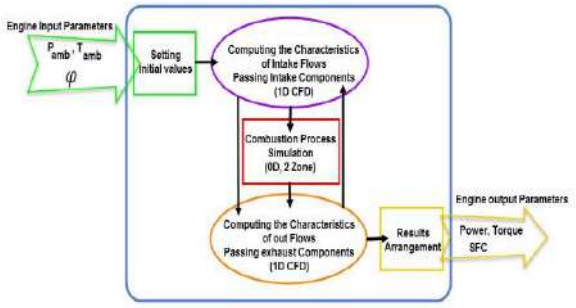


Figure 2. Schematic of one-dimensional component model with a zero-dimensional combustion sub-model

2. 1. Component Simulation One-dimensional simulation needs to solve mass, energy and momentum equations, simultaneously. The trapped mass of each component is calculated by mass conservation equation [32]. The energy equation is expanded in transient form using work and convectional heat loss definitions [33] and momentum conservation equation is considered along with each component [32].

$$\dot{m}_{sub} = \sum \dot{m}_e - \sum \dot{m}_i \quad (1)$$

$$\frac{d(me)}{dt} = P\frac{dV}{dt} + \sum_i \dot{m}_i h_i - \sum_e \dot{m}_e h_e - h_g A(T_{gas} - T_{wall}) \quad (2)$$

$$\frac{\dot{m}}{dt} = \frac{dpA + \sum_i \dot{m}_i u + \sum_e \dot{m}_e u}{dx} - \frac{4 C_f \frac{\rho u^2}{2} \frac{dxA}{D} - C_p (\frac{1}{2}\rho u^2)A}{dx} \quad (3)$$

In these equations, $\dot{m}_i$ and $\dot{m}_e$ are the mass flow rates at the entrance and exit of each component which are defined as,

$$\dot{m} = \rho U A \quad (4)$$

where ρ, A, U indicate density, area and velocity of flow. Indeed, P, V, e, h, h_g, T_{gas} and T_{wall} in the energy equation are pressure, volume, specific internal energy, specific enthalpy, convection heat transfer coefficient, flow temperature and wall temperature, respectively. Convection heat transfer coefficient is described as:

$$h_g = \frac{1}{2} \rho C_f U_{eff} C_p Pr^{-\frac{2}{3}} \quad (5)$$

Here, C_f , U_{eff}, C_p and Pr define friction coefficient, effective velocity, specific heat coefficient and Prandtl number, respectively. Considering pipes roughness, the friction coefficient is defined by Nikuradse equation [34],

$$C_{f\,(rough)} = \frac{0.25}{(2\log_{10}(\frac{1D}{2h})+1.74)^{0.25}} \quad (6)$$

where, D is the diameter of pipe and h is the height of roughness. In the momentum equation, C_p is pressure loss coefficient which is defined as:

$$C_p = \frac{P_i - P_e}{\frac{1}{2}\rho U_i^{\,2}} \quad (7)$$

Indexes i and e show inlet and outlet conditions.

2. 2. Combustion Simulation A zero-dimensional two-zone thermodynamic model is employed for combustion process simulation. The combustion chamber is divided into the burned and unburned zones and the first law of thermodynamics, ideal gas equation of state besides engine geometrical correlations are applied on each zone. The schematic of the two-zone model is shown in Figure 3. It should be noted that the heat transferred between two zones is ignored and it is assumed that a proportion of charge, due to the Wiebe function, enters the burned zone in each time-step.

Heat release rate is also defined by the Wiebe function which is modified for gasoline blend combustion [35]

$$x_b = 1 - \exp\left(-Ea\left(\frac{\theta-\theta_{ig}}{\Delta\theta}\right)^{m+1}\right) \quad (8)$$

where, x_b, Ea, and θ_{ig} are the mass fraction of burnt fuel, activation energy, and spark time, respectively. Energy equation, assuming charge and combustion products as the ideal gases and considering SI combustion process, can be written as [35]:

$$\frac{d(M_B E_B)}{d\theta} + \mathrm{P}\frac{\mathrm{dV_B}}{d\theta} + \mathrm{Q_B} = \eta\ \mathrm{LHV}\ \mathrm{M_f}\frac{dx_b}{d\theta} + h_U\frac{dM_B}{d\theta} \quad (9)$$

$$\frac{d(M_U E_U)}{d\theta} + \mathrm{P}\frac{\mathrm{dV_U}}{d\theta} + \mathrm{Q_U} = h_U\frac{dM_U}{d\theta} \quad (10)$$

where, M, E, V, P, η, LHV , and h are the symbols of mass, internal energy, volume, pressure, combustion efficiency, the low heating value of fuel, and specific enthalpy and the indexed U and B refer to the unburnt and burnt zones. The heat transferred from the unburnt zone is equal to $x_b Q$. It is considered as $(1-x_b)Q$ for the burnt zone, where Q is the total heat transfer from the Woschni correlation [36].

$$\frac{d\mathrm{Q}}{d\theta} = h_c A_c \frac{dT}{d\theta} \quad (11)$$

$$h_c = 130\ P^{0.8}U^{0.8}B^{-0.2}T^{-0.55} \quad (12)$$

Here, A_C, T, and U are the effective area of heat transfer, in-cylinder gas temperature, and gas local speed which is calculated by the mean piston velocity. It should be noted that the temperature of these equations are the mean cylinder temperature which is calculated as [35]:

$$T = \frac{x_b C_{v_B} T_B + (1-x_b)C_{v_U}T_U}{x_b C_{v_B} + (1-x_b)C_{v_U}} \quad (13)$$

C_v is the specific heat coefficient in constant volume and the volume of combustion chamber defined by engine geometrical correlation [35].

$$V = V_c + \frac{\pi B^2}{4}\left(l - a - a\cos(\theta) - \sqrt{l^2 - a^2\sin^2(\theta)}\right) \quad (14)$$

Here, $\mathrm{V_c}$, l, and a are the clearance volume, connecting rod length and crank radios, respectively and θ refers to the crank angle. The total volume is divided into the two burnt and unburnt sections, so the state equation of ideal gas for each zone is like;

$$\frac{PV_B}{RT_B} = M_B = x_b M_t \quad (15)$$

$$\frac{PV_U}{RT_U} = M_U = (1-x_b)M_t \quad (16)$$

where, M_t refers to the total trapped mass in the cylinder. In addition to the noted correlations, there are some other considerations applied to the control volume due to the component type, its characteristics and operating conditions. Some of them are listed in the following;

- The velocity, turbulence, pressure loss and other features of charge flow would be affected by the valve lifting profile.
- The Continuous Variable Valve Timing (CVVT) is adopted from the engine operating map.
- The losses caused by the injection type would be affected by the characteristics of the injector such as static injection, dynamic injection, injection duration, injection timing, injection angle, and the number of nuzzles. In this study, only the effect of the number of nuzzles is ignored.
- The turbocharger is modelled by the own operating map which is provided by the producer, so the turbine rotational speed, efficiency, power and the boosted pressure are calculated by the amount of flow passing the turbine blades.
- The passing flow of turbine blades is controlled by the wastegate lifting profile which is adopted from the engine operating map.
- Ignition timing, fuel injected, and equivalence ratio are also adopted from the engine operating map.
- The friction loss of engine components is defined due to the engine frictional test.
- The compressor bypass flow, cyclic variations, angle of throttle, added fuel by canister valve, the losses by oil pump, water pump, alternator and other accessories are ignored due to the kind of simulation.

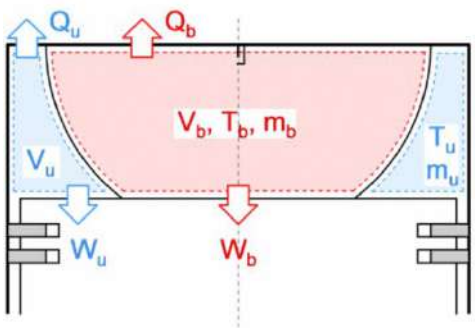


Figure 3. Schematic of two-zone combustion model

3. EXPERIMENTAL SETUP

The experimental tests were carried out in the Irankhodro Powertrain Company (IPCO). Two types of Iranian national engines, EF7, were available for the experimental test: Natural Aspirated (NA) and Turbocharged (TC). The main characteristics of the EF7 engines are reported in Table 2.

The schematic of the test room is shown in Figure 4 and also the accuracy of measurement instruments is reported in Table 3. Both of the engines are coupled with the AVL 220kW dynamometer which is controlled by the Puma controling system and able to fix the engine speed by 1rpm accuracy. The utilized fuel flow meter, fuel temperature control system and blowby meter are licenced by the AVL company. Indeed, the combustion process is analyzed separately in the Indicom commercial sowftware environment and the engine actuators are generally controled by the other computer; thanks to employing open ECU and INCA commercial software. The used Horiba gas analyzer brings this opportunity that investigates the inlet air characteristics and measuring the equivalance ratio, in addition to evaluate the concentration of the exhaust gas species.

4. VALIDATION

To investigate the EF7 engine performance and also the capability of its downsizing, two one-dimensional

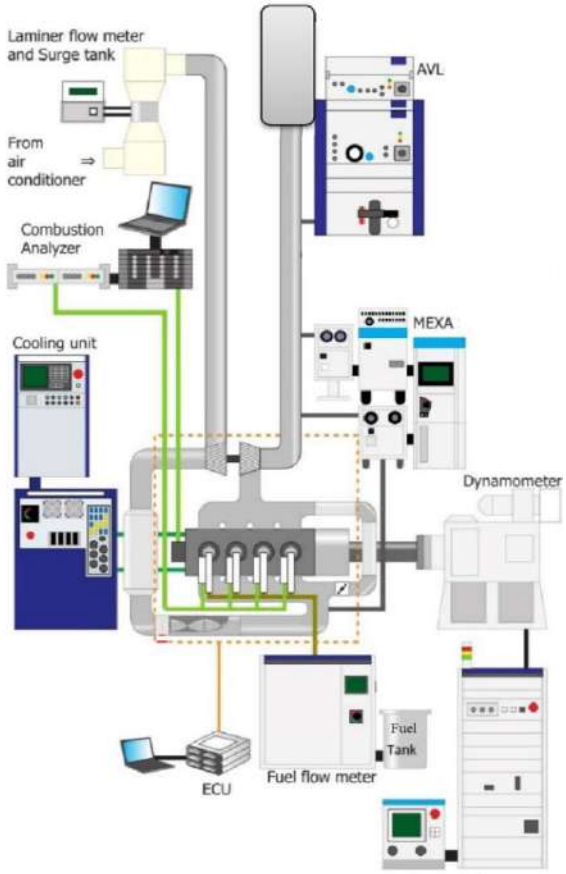


Figure 4. Test setup

TABLE 2. The main characteristics of EF7 engines

Engine name	EF7NA [37]	EF7TC [38]
Engine type	4 inline cylinder	
Bore × Stroke	78.6 × 85 mm	
Connection rod length	134.5 mm	
Compression Ratio	11	9.6
IVC	40 deg aBDC	26 deg aBDC
Exhaust Valve Opening (EVO)	50 deg bBDC	25 deg bBDC
Fuel	Gasoline	
CVVT	Intake	---

TABLE 3. Specifications of measurement instruments

Parameter	Accuracy
In-cylinder Pressure	±0.1 bar
Engine Speed	±1 rpm
Crank Angle	±0.1 CAD
Torque	±0.1 N.m
Fuel Consumption	0.12%
Equivalence Ratio	±0.02

models are provided using described equations at model description section within the GT-Power commercial software environment. The first model simulates the NA engine and the second one is developed to simulate TC engine performance.

The results of simulations for NA and TC engines at full load condition are compared with the experimental data. It should be noted that, all engine input parameters are adopted from the engine operating map. Brake torque, power, Mean Effective Pressure (BMEP), Specific Fuel Consumption (BSFC) in different engine speeds are considered as the main parameters for validation. The results for the NA engine showed the mean error of brake torque by 5.44%, specific fuel consumption by 2.74% and brake power by 4.22% illustrated in Figures 5 to 7. Simulated brake power, torque and also in-cylinder pressure variation are also compared via experimental results for TC engine shown in Figures 8 to 10; which show the meanly error of 2.36% for brake power, 2.93% for brake torque and acceptable estimation of in-cylinder pressure. It should be noted that the experimental result for in-cylinder pressure is shown for the 500-cycles sample.

Considering Figures 5 to 10, it would be asserted that the provided models are reliable for both EF7NA and TC

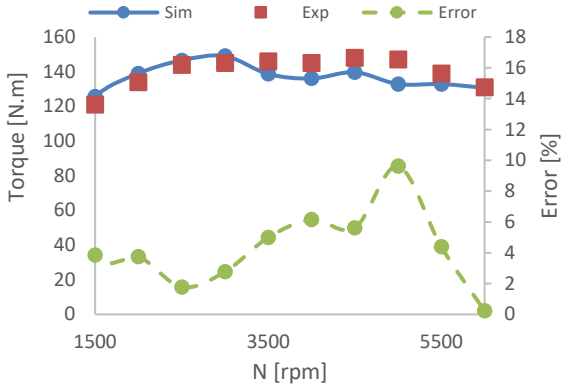


Figure 5. Simulated brake torque via experimental results for EF7NA

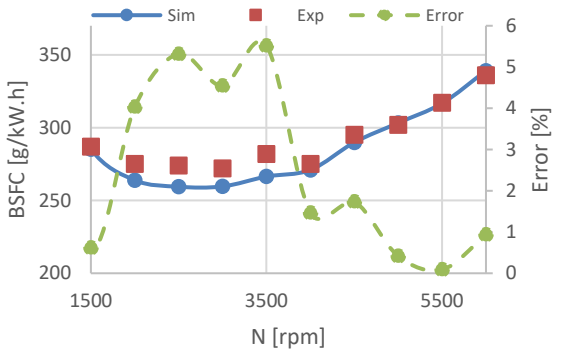


Figure 6. Simulated brake specific fuel consumption via experimental results for EF7NA

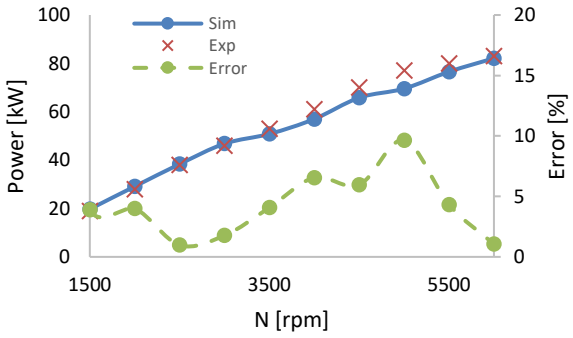


Figure 7. Simulated brake power via experimental results for EF7NA

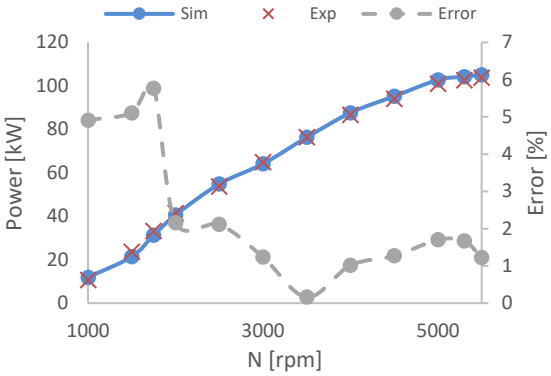


Figure 8. Simulated brake power via experimental results for EF7TC

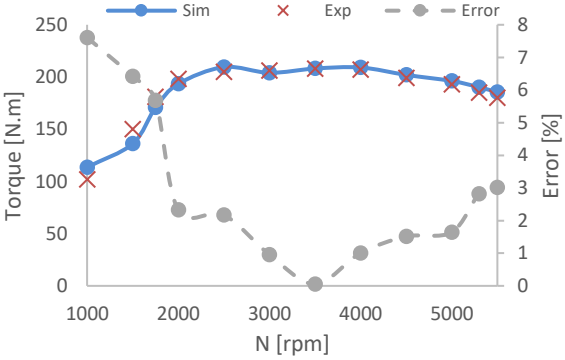


Figure 9. Simulated brake torque via experimental results for EF7TC

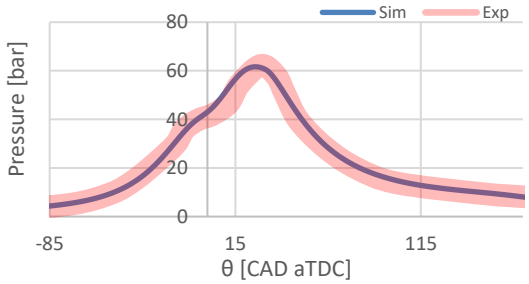


Figure 10. Simulated in-cylinder pressure via experimental results for EF7TC, N=5500 rpm

engines performance evaluation, and also it can be used for feasibility study of downsizing.

5. RESULTS AND DISCUSSION

Gasoline fueled EF7NA engine is considered as the base engine for downsizing. The main objective of this study is to propose a downsized EF7 version employing the technology available in Iran. At first, considering the gasoline fueled EF7TC engine performance, a 3-cylinder gasoline fueled TC engine is introduced as the first conceptual version of downsized EF7 called EF7α. In the next step, the role of employing different turbochargers is investigated to improve the performance of EF7α. Finally, to achieve acceptable performance in low speed conditions, the effect of employing VVT is also studied. The main characteristics of the base engine besides EF7α are reported in Table 4.

It should be noted that, in this step, the fuel cut-off strategy for a cylinder of EF7TC, is applied at the test setup to have an estimation of experimental performance of the 3-cylinder engine. In this study, the results of this strategy are called cylinder deactivated. However, the simulations are carried out for both strategies; cylinder deactivated and real 3-cylinder mode.

The concept of EF7α is adopted by gasoline fueled EF7TC engine performance which is compared with NA one in Figure 11. As it is shown in Figure 11, the TC engine produces brake torque almost 1.33 times more than NA one. Consequently, it can be expected that EF7α shows the same performance as the considered base engine. The results of simulations for both strategies; cylinder deactivated and real 3-cylinder mode are compared with the base engine shown in Figure 12. The results showed the well agreement with the noted assertion in case of cylinder deactivated and poor performance for 3-cylinder mode before 2800 rpm. The main reason of such a behavior is due to the noticeable reduction in the flow-rate passing turbine blades in case of 3-cylinder mode. However setting the right situation of the wastegate (considering the limitations such as knock, maximum boost pressure, maximum turbine inlet temperature and maximum turbine speed), the EF7α provides the same torque as the base engine after 2800 rpm. The same challenges had been reported in the litrature and different sterategies were employed to tacle them [24, 26, 28].

TABLE 4. The basic data of base and downsized engines

Engine name	Base Engine	EF7α
Engine type	4 inline cylinder	3 inline cylinder
	NA	TC
Bore × Stroke	78.6 × 85 mm	
Connection rod length	134.5 mm	
Compression ratio	11	9.6
Fuel	gasoline	
CVVT	Intake	Intake, Exhaust

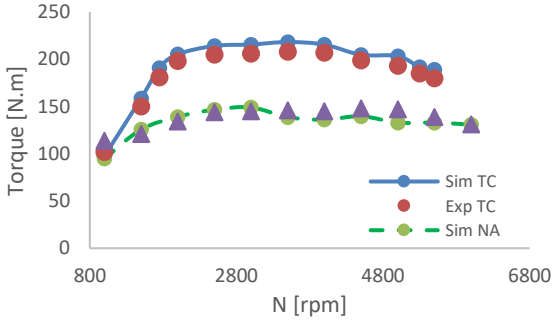


Figure 11. Brake torque comparison, EF7TC via EF7NA

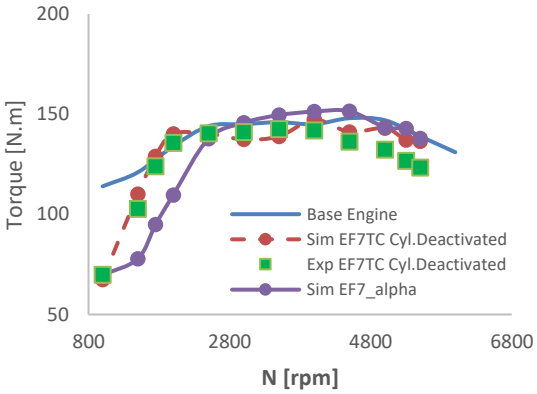


Figure 12. Brake torque comparison, EF7α via base engine

Due to the Figure 12, it is obvious that to achieve better performance at low speed conditions, employing the turbocharger well-matched with EF7α, is needed. To achieve this aim, a turbo-matching study is carried out based on the common turbochargers of Iran markets which are able to be installed on EF7. The basic information of investigated turbochargers is reported in Table 5. The operating maps of these turbochargers are applied to the model and the results are compared in Figure 13. Looking more detail in Figure 13, it can be asserted that from the investigated turbochargers, the F-Diesel one has the best performance matching with EF7α which shifts the engine acceptable performance from >2800 rpm to >1900 rpm.

Although using the F-Diesel turbocharger the performance of EF7α seems a better situation, the performance enhancement between 1000 to 1900 rpm is still needed. To achieve this aim, employing VVT technology can be considered as a promising solution. Consequently, the VVT sweep is carried out before 2000 rpm thanks to the ability of provided numerical model. In this approach dual CVVT is applied to the model and finally the best valve timing for each speed selected based on the maximum torque and minimum BSFC. The swept CVVTs at 1500 rpm are reported in Table 6. The candidate positions have shaded table cells and the desired CVVT are underlined in the table. The engine

TABLE 5. The information of investigated turbochargers

Case	Company	Model	Base Installed Engine
TC#1	Fuyuan	JEF7-3B	EF7TC, Commercial (Gasoline Fueled)
TC#2	Honeywell	C44-T79	EF7TC, Mule (Bi-Fueled)
TC#3		C46-T79	EF7TC, Mule (CNG Dedicated)
TC#4	F-Diesel	FD159-A016	EF7TC+, Mule (Gasoline Fueled)

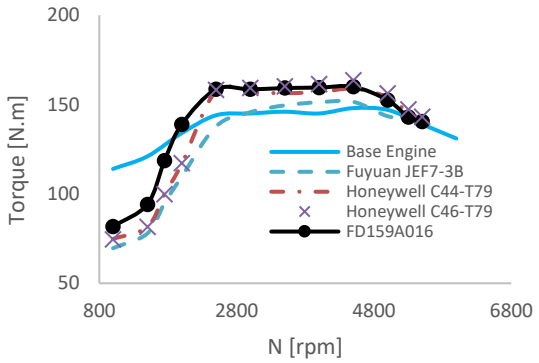


Figure 13. Brake torque comparison, EF7α with deferent turbochargers via base engine

behaivour shown in Table 6 is affected by different controling/limiting parameteres such as volumetric efficiency, real compression ratio, knock, spark timing, turbine inlet maximum temperature, turbine maximum rotational speed, maximum pressure/temperature in inlet port and allowed range of equvalance ratio. Although, each of these parameters may be the limitting issue to achive more torque in a certain case, but in general, the main cause of such a behavoiur is the variations of volumetric efficiency and real compression ratio due to the valve overlap increase. As an expample the trend of volumetric efficiecy variation due to the change in CVVT shown in Figure 14, is the same of reported torque in Table 6. However, gained torque is affected by the other noted parameters simultaneously. The results of employing desired CVVT at the lwo speed conditions are shown in Figure 15, and it is clear that the provided torque of EF7α using optimum VVT and well-matched turbocharger is at least the same as the base engine. However, the provided torque at 1000 rpm is still estimated 7.01% lower than the base engine which is acceptable for a downsized engine [21].

The comparison of brake power is also shown the same trend for EF7α, and the engine is able to provide at least the same power as the base engine after 1500 rpm. However, the provided power at 1000 rpm is still estimated 4.5 kW lower than the base engine, as shown in Figure 16.

BMEP and BSFC of EF7α are also compared with the base engine, shown in Figures 17 and 18. It is obvious in both figures that the performance of EF7α in real 3-cylinder mode covers the base engine while there are significant differences for cylinder deactivated approach (32.5% decrease in BMEP and 57% increase in BSFC). These are related to the definition of these parameters that

TABLE 6. The swept CVVTs, and definition of desired VVT at 1500 rpm

	Exhaust=0		Exhaust=-17		Exhaust=-33		Exhaust=-52	
	Torque [N.m]	**BSFC [g/kW.h]**	**Torque [N.m]**	**BSFC [g/kW.h]**	**Torque [N.m]**	**BSFC [g/kW.h]**	**Torque [N.m]**	**BSFC [g/kW.h]**
Intake =0	100.8	246.5	100.9	227.4	92.2	228.4	80.0	236.3
Intake =15	115.5	267.8	**117.7**	**259.4**	105.7	258.2	119.9	271.6
Intake =30	109.0	247.7	**116.7**	**259.8**	136.8	277.7	142.2	354.7
Intake =45	107.4	260.7	127.7	276.4	144.3	335.4	88.5	376.0
Intake =60	98.3	238.1	124.4	279.3	98.1	355.7	75.1	302.0

*CVVT Intake=0, Exhaust=0 is referred to the lock position which means no overlap for valve timing

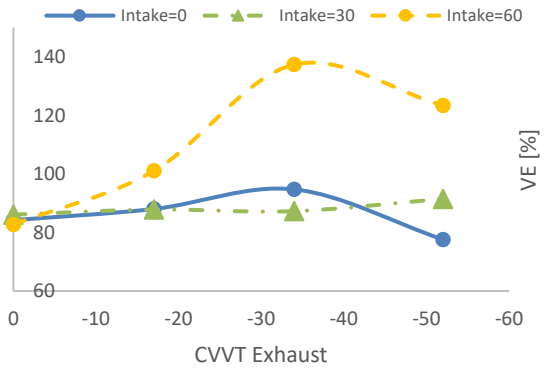


Figure 14. Volumetric efficiency variation via CVVT

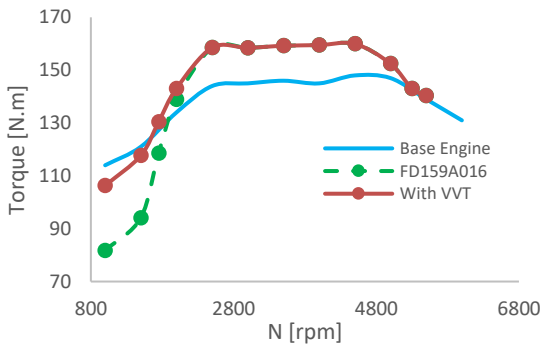


Figure 15. Brake torque comparison, EF7α with turbocharger and CVVT via base engine

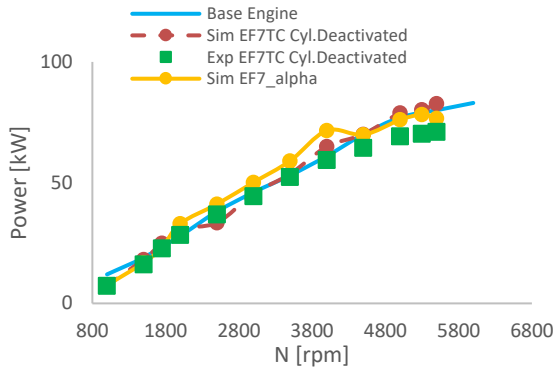


Figure 16. Brake power comparison, EF7α via base engine

the reduction in fuel consumption is occurred just for a cylinder while the power loss increased by the frictional and pumping losses of the deactivated cylinder. Further than the lack of pumping loss of the deactivated cylinder, the frictional and inertial resistances due to the engine components such as camshaft are also removed in real 3-cylinder mode. Similarly, the effect of distance-volume reduction is not applied in cylinder deactivated approach.

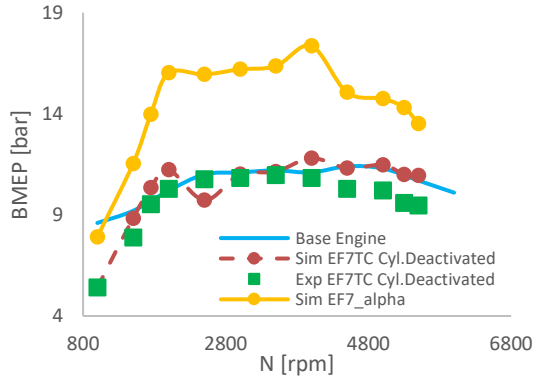


Figure 17. BMEP comparison, EF7α via base engine

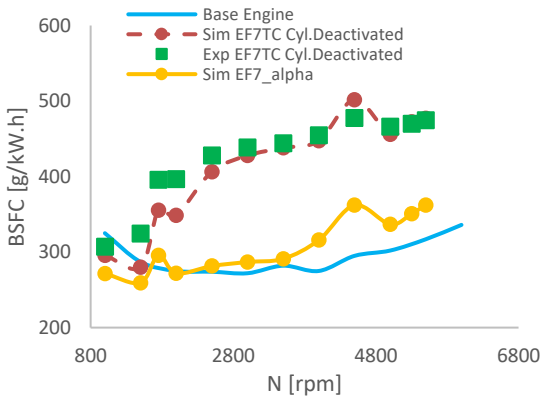


Figure 18. BSFC comparison, EF7α via base engine

6. CONCLUSION

In this research the Iranian engine EF7NA is evaluated for 25% downsizing. Cylinder reduction besides using boost technologies namely; turbocharger and CVVT is the main employed strategy for this research. To investigate the downsized engine performance, two one-dimensional models equipped with a two-zone combustion sub-model and an experimental test cell are employed which are able to apply different strategies of this work. The main results are listed as follow:

- Provided models have enough accuracy to investigate the full load operation of both; NA and TC engines.
- Provided models are able to estimate the effect of using CVVT and different turbochargers.
- Both strategies; cylinder deactivation and cylinder reduction are applicable in the provided models.
- The 3-cylinder strategy employing F-Diesel as the well-matched turbocharger and CVVT can cover all the requirements as the downsized engine.
- EF7α can be introduced as the first generation of Iranian downsized engine.

7. AKNOWLEDGMENT

The authors wish to express their appreciation to the Irankhodro Powertrain Company (IPCO) for sharing necessary data during the course of this research.

8. REFERENCES

1. International Energy Agency. Policy pathways: improving the fuel economy of road vehicles – A policy package; 2011. https://www.iea.org/publications/freepublications/publication/policy-pathways-improving-the-fuel-economy-of-road-vehicles---a-policy-package.html
2. Zhang, W., Yang, X., Wang, T., Peng, X. and Wang, X., "Experimental study of a gas engine-driven heat pump system for space heating and cooling", ***Civil Engineering Journal***, Vol. 5, No. 10, (2019), 2282-2295. https://doi.org/10.28991/cej-2019-03091411
3. Borowski, P.F., "New technologies and innovative solutions in the development strategies of energy enterprises", ***High Tech and Innovation Journal***, Vol. 1, No. 2, (2020), 39-58. https://doi.org/10.28991/HIJ-2020-01-02-01
4. Topçuoğlu, K., "Trombe wall application with heat storage tank", ***Civil Engineering Journal***, Vol. 5, No. 7, (2019), 1477-1489. https://doi.org/10.28991/cej-2019-03091346
5. Ravi, P., Devanandh, V., Pandey, S.K., Senthilnathan, K., Sadagopan, K. and Patel, B.P., Quasi-dimensional thermodynamic simulation study of downsizing on a four-cylinder turbocharged engine, ***in Advances in Energy Research***, vol. 2, (2020), 563-575. https://doi.org/10.1007/978-981-15-2662-6_51
6. Namar, M.M., Jahanian, O., Shafaghat, R. and Nikzadfar, K., "Feasibility study for downsizing ef7 engine, numerical and experimental approach", ***The Journal of Engine Research***, Vol.

61, No. 61, (2021), 73-85. http://engineresearch.ir/article-1-758-en.html

7. Cho, J., Kim, K., Baek, S., Myung, C.-L. and Park, S., "Abatement potential analysis on CO_2 and size-resolved particle emissions from a downsized lpg direct injection engine for passenger car", ***Atmospheric Pollution Research***, Vol. 10, No. 6, (2019), 1711-1722. https://doi.org/10.1016/j.apr.2019.07.002
8. Knauder, C., Allmaier, H., Sander, D.E. and Sams, T., "Investigations of the friction losses of different engine concepts. Part 2: Sub-assembly resolved friction loss comparison of three engines", ***Lubricants***, Vol. 7, No. 12, (2019), 105. https://doi.org/10.3390/lubricants7120105
9. Namar, M.M. and Jahanian, O., "Energy and exergy analysis of a hydrogen-fueled hcci engine", ***Journal of Thermal Analysis and Calorimetry***, Vol. 137, No. 1, (2019), 205-215. https://doi.org/10.1007/s10973-018-7910-7
10. Jafari, B., Khatamnejad, H., Shahavi, M.H. and Domeyri Ganji, D., "Simulation of dual fuel combustion of direct injection engine with variable natural gas premixed ratio", ***International Journal of Engineering, Transactions C: Aspects***, Vol. 32, No. 9, (2019), 1327-1336. https://dx.doi.org/10.5829/ije.2019.32.09c.14
11. Namar, M.M. and Jahanian, O., "A simple algebraic model for predicting hcci auto-ignition timing according to control oriented models requirements", ***Energy Conversion and Management***, Vol. 154, (2017), 38-45. https://doi.org/10.1016/j.enconman.2017.10.056
12. Hassanzadeh Saraei, S., Jafarmadar, S., Khalilarya, S. and Taghavifar, H., "Effects of triple injection strategies on performance and pollutant emissions of a di diesel engine using cfd simulation", ***International Journal of Engineering, Transactions C: Aspects***, Vol. 31, No. 6, (2018), 973-979. https://dx.doi.org/10.5829/ije.2018.31.06c.15
13. Kazemian, M. and Gandjalikhan Nassab, S., "Thermodynamic analysis and statistical investigation of effective parameters for gas turbine cycle using the response surface methodology", ***International Journal of Engineering, Transactions B: Applications***, Vol. 33, No. 5, (2020), 894-905. https://dx.doi.org/10.5829/ije.2020.33.05b.22
14. Hu, K. and Chen, Y., "Technological growth of fuel efficiency in european automobile market 1975–2015", ***Energy Policy***, Vol. 98, (2016), 142-148. https://doi.org/10.1016/j.enpol.2016.08.024
15. Merker, G.P., Schwarz, C. and Teichmann, R., "Combustion engines development: Mixture formation, combustion, emissions and simulation, ***Springer Science & Business Media***, (2011). https://doi.org/10.1007/978-3-642-14094-5
16. Kuhlbach, K., Mehring, J., Borrmann, D. and Friedfeld, R., "Zylinderkopf mit integriertem abgaskrümmer für downsizing-konzepte", ***MTZ-Motortechnische Zeitschrift***, Vol. 70, No. 4, (2009), 286-293. http://dx.doi.org/10.1007/BF03225480
17. Smith, A., "Stroke of genius for gasoline downsizing", ***Ricardo Q Review***, (2008), https://scholar.google.com/scholar_lookup?title=Stroke+of+genius+for+gasoline+downsizing&publication_year=2008
18. Budack, R., Wurms, R., Mendl, G. and Heiduk, T., "Der neue 2, 0-l-r4-tfsi-motor von audi", MTZ-Motortechnische Zeitschrift, Vol. 77, No. 5, (2016), 16-25. http://dx.doi.org/10.1007/s35146-016-0035-2
19. Wang, Y., Wei, H., Zhou, L., Li, Y. and Liang, J., "Effect of injection strategy on the combustion and knock in a downsized gasoline engine with large eddy simulation", (2020), ***SAE Technical Paper***. https://doi.org/10.4271/2020-01-0244
20. Morikawa, K., Shen, F., Yamada, T., Moriyoshi, Y. and Kuboyama, T., "The extension of load range and low fuel consumption range based on the ultra-highly boosted downsized engine concept", ***Transactions of Society of Automotive Engineers of Japan***, Vol. 51, No. 5, (2020). https://doi.org/10.11351/jsaeronbun.51.862
21. Hancock, D., Fraser, N., Jeremy, M., Sykes, R. and Blaxill, H., "A new 3 cylinder 1.2 l advanced downsizing technology demonstrator engine" (2008), ***SAE Technical Paper***. https://doi.org/10.4271/2008-01-0611
22. Dönitz, C., Vasile, I., Onder, C. and Guzzella, L., "Realizing a concept for high efficiency and excellent driveability: The downsized and supercharged hybrid pneumatic engine", (2009), ***SAE Technical Paper***, https://doi.org/10.4271/2009-01-1326
23. Splitter, D.A. and Szybist, J.P., "Experimental investigation of spark-ignited combustion with high-octane biofuels and egr. 1. Engine load range and downsize downspeed opportunity", ***Energy & Fuels***, Vol. 28, No. 2, (2014), 1418-1431. https://doi.org/10.1021/ef401574p
24. Attar, M.A., Herfatmanesh, M.R., Zhao, H. and Cairns, A., "Experimental investigation of direct injection charge cooling in optical gdi engine using tracer-based plif technique", ***Experimental Thermal and Fluid Science***, Vol. 59, (2014), 96-108. http://dx.doi.org/10.1016/j.expthermflusci.2014.07.020
25. Turner, J., Popplewell, A., Patel, R., Johnson, T., Darnton, N., Richardson, S., Bredda, S., Tudor, R., Bithell, C. and Jackson, R., "Ultra boost for economy: Extending the limits of extreme engine downsizing", ***SAE International Journal of Engines***, Vol. 7, No. 1, (2014), 387-417. https://doi.org/10.4271/2014-01-1185
26. Millo, F., Luisi, S., Borean, F. and Stroppiana, A., "Numerical and experimental investigation on combustion characteristics of a spark ignition engine with an early intake valve closing load control", ***Fuel***, Vol. 121, (2014), 298-310. https://doi.org/10.1016/j.fuel.2013.12.047
27. Severi, E., d'Adamo, A., Berni, F., Breda, S., Lugli, M. and Mattarelli, E., "Numerical investigation on the effects of bore reduction in a high performance turbocharged gdi engine. 3d investigation of knock tendency", ***Energy Procedia***, Vol. 81, No., (2015), 846-855. https://doi.org/10.1016/j.egypro.2015.12.094
28. Ma, J. and Zhao, H., "The modeling and design of a boosted uniflow scavenged direct injection gasoline (busdig) engine" (2015), ***SAE Technical Paper***. https://doi.org/10.4271/2015-01-1970
29. Wróblewski, E., Finke, S. and Babiak, M., "Investigation of friction loss in internal combustion engine of experimental microgeometry piston bearing surface", ***Journal of KONES***, Vol. 24, (2017). https://doi.org/10.5604/01.3001.0010.2951
30. Liu, J., Liu, Y. and Bolton, J.S., "The application of acoustic radiation modes to engine oil pan design" ***SAE Technical Paper***, (2017), https://doi.org/10.4271/2017-01-1844
31. Gamma Inc. Users Maunual, volume 61. Gamma Technologies, 2004. https://www.gtisoft.com/
32. White, F.M., Fluid mechanics, WCB. Ed McGraw-Hill Boston; 1999. https://books.google.com/books?id=fa_pAAAAMAAJ
33. Sonntag RE, Borgnakke C, Van Wylen GJ, Van Wyk S. Fundamentals of thermodynamics. New York: Wiley; 1998. https://books.google.com/books?id=95hVAAAAMAAJ
34. Nikuradse J. "Laws of flow in rough pipes", Washington: National Advisory Committee for Aeronautics; 1950 Nov 1. https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.467.2980&rep=rep1&type=pdf
35. Heywood JB. Internal Combustion Engine Fundamentals. 1st Edition, McGraw-Hill, 1988. https://books.google.com/books?id=O69nQgAACAAJ
36. Woschni, G., "A universally applicable equation for the instantaneous heat transfer coefficient in the internal combustion engine" (1967), ***SAE Technical Paper***, https://doi.org/10.4271/670931

37. Asgari, O., Hannani, S.K. and Ebrahimi, R., "Improvement and experimental validation of a multi-zone model for combustion and no emissions in cng fueled spark ignition engine", ***Journal of Mechanical Science and Technology***, Vol. 26, No. 4, (2012), 1205-1212. https://doi.org/10.1007/s12206-012-0229-6

38. Gharloghy J, Kakaee A, Forooghifar A. Comparison of EF7 TC engine performance in two modes of CNG and petrol using piston and combustion chamber thermal simulation. ***Fuel and Combustion***, (2011), Vol. 3, No. 2, 1-16. https://www.sid.ir/en/journal/ViewPaper.aspx?ID=197151

Persian Abstract

چکیده

کوچک سازی ابعادی به عنوان یک روش پذیرفته شده برای کاهش سوخت مصرفی و آلایندگی موتورهای احتراق داخلی مطرح می‌باشد. در این پژوهش موتور ملی تفس طبیعی EF7 به عنوان هدف کوچک سازی ابعادی انتخاب شده و در راستای ارائه یک نسخه ابعاد کوچک از آن تلاش شده است. بدین منظور از یک مدل عددی یک بعدی موتور استفاده شده که قابلیت اعمال استراتژی‌های مختلف شامل افزودن توربوشارژر، غیرفعال سازی یک سیلندر، حذف یک سیلندر و همچنین زمانبندی متغییر سوپاپ‌ها را دارد. به علاوه، به منظور صحت سنجی نتایج مدل، از یک بستر آزمایشگاهی جهت بررسی عملکرد موتور در حالت‌های تنفس طبیعی، توربوشارژ و یک سیلندر غیر فعال شده استفاده شده است. به منظور جبران ضعف عملکردی موتور ابعاد کوچک در دورهای پایین، تحلیل عملکرد موتور در ازای بکارگیری توربوشارژرهای مختلف و همچنین زمانبندی مختلف سوپاپ‌ها صورت پذیرفته و نتایج بدست آمده حاکی از آن است که موتور ابعاد کوچک معرفی شده، با بکارگیری توربوشارژر و زمانبندی بهینه سوپاپ‌ها می‌تواند گشتاوری معادل موتور پایه را ارئه دهد. حداکثر اختلاف گشتاور ارائه شده بربر با ۷ درصد و در دور ۱۰۰۰ دور بر دقیقه می‌باشد.

IJE TRANSACTIONS C: Aspects, Vol. 34, No. 09, (September 2021) 2148-2156

International Journal of Engineering

Journal Homepage: www.ije.ir

Adaptive Neuro-Fuzzy Inference System Estimation Propofol Dose in Induction Phase During Anesthesia: A Case Study

N. Jamali[a], A. Sadegheih*[a], M. M. Lotfi[a], H. Razavi[b]

[a] *Faculty of Industrial Engineering, Yazd University, Yazd, Iran*
[b] *Department of Industrial Engineering, Faculty of Engineering, Ferdowsi University of Mashhad, Mashhad, Iran*

PAPER INFO

Paper history:
Received 14 November 2020
Received in revised form 13 March 2021
Accepted 28 July 2021

Keywords:
Propofol dose
Anesthesia
ANFIS
Estimation Model
Intravenous Anesthetic
Induction

ABSTRACT

In this study, the anesthetic drug dose is estimated with respect to patients' physiological parameters. The most critical anesthetic drug, propofol is considered in this modeling. The intravenous propofol is one of the widely used for both induction and maintenance phases of anesthesia. According to a deep uncertainty estimation model, the adaptive neuro-fuzzy inference system is applied to estimate the safe dose of anesthetic drug. The propofol model is estimated based on the patients' physiological parameters (age, weight, height, and gender) and variables (blood pressure, heart rate, and depth of anesthesia) each time. The sensitivity analysis evaluates the validity of the estimation model. At the end, performance of the proposed estimation model is compared to that of the classical Pharmacokinetics-Pharmacodynamics (PK-PD) model and the data obtained from the patients undergoing surgery. The results showed that the Adaptive Neuro-Fuzzy Inference System (ANFIS) estimation model with accuracy of 0.999 reduces the total amount of propofol dose. The novelty of the proposed model in this paper lies in its estimation of the depth of anesthesia in induction separately from the maintenance phase independent of Bispectral Index (BIS). To validate our methodology, a real case study of Mashhad hospital in Iran has provided, resulting in a comprehensive discussion and managerial insights.

***doi**: 10.5829/ije.2021.34.09c.12*

NOMENCLATURE

DOA	Depth of Anesthesia	C_1	Drug concentration
ANFIS	Adaptive Neuro-Fuzzy Inference System	C*2*	Slow compartments drug concentrations
BIS	Bispectral Index	C*3*	Fast compartments drug concentrations
BP	Blood Presure	kij	Drug transfer from ith to jth compartment
HR	Heart Rate	$\dot{C}_e(t)$	Centeral compartment of effect site
PK	Pharmacokinetics	lbm	Lean Body Mass
PD	Pharmacodynamics	**Greek Symbols**	
$\mu_{A_j^i}$	Membership Function(MF)	Γ	curve slope

1. INTRODUCTION

The general anesthesia complexity process consists of analgesia, hypnosis, and neuromuscular blockade. The anesthesia administration is traditionally carried out by anesthetists, who, according to their experience and personal knowledge, decide on the dose of adequately initial bolus as well as the dose of anesthetic drug to maintain a proper relaxation level during the surgery. A basic requirement for a safe anesthetic is to make decision on the proper combination of the type and dose of drugs under the direct supervision of the anesthesiologist [1]. Propofol is considered as the most popular intravenous anesthetic agent having short duration of action and rapid onset in the induction and

*Corresponding Author: sadegheih@yazd.ac.ir (A. Sadegheih)

Please cite this article as: N. Jamali, A. Sadegheih, M. M. Lotfi, H. Razavi, Adaptive Neuro-Fuzzy Inference System Estimation Propofol Dose in Induction Phase During Anesthesia: A Case Study, International Journal of Engineering, Transactions C: Aspects, Vol. 34, No. 09, (2021) 2148-2156

maintenance during the general anesthesia process [2]. However, the prediction accuracy models may be affected by the patients' parameters. A current review of pharmacokinetics and pharmacodynamics approaches to optimize the anesthetic drug administration was performed [3, 4]. In this pursuit, advanced mathematical techniques can be of superb value, for instance, robust PID control [5], fuzzy-type I/II [6, 7], propofol injection robot-assistant [8, 9] and target control infusion [10]. Some of the auto-drug delivery systems have widely used bispectral index monitoring [11, 12], as variables of control modeling to administer the propofol dose during the anesthesia. Although there are some limitations such as inadequate control in the induction phase, lack of the pysiological parameters of the patients anesthetized, and the variability of the patient situation during the model-based controllers process. Machine learning and artificial intelligence techniques are also highly applicable and suitable to modify prediction and estimation models by using real-time observed data from patients. Adaptive Neuro-Fuzzy Inference System (ANFIS), widely applied for estimation in many works [13-15].

It should be noted that auto-control anesthesia model and auto-injection system are depending on the patients' depth of anesthesia. Bispectral Index (BIS) is a dimensionless indicator to show the patients' depth of anesthesia during surgery [16]. Having an accurate model in the predictive controller is the main problem for the process under control. The prediction accuracy mainly depends on the parameters of the model. This study contributes to the estimation of the propofol dose according to the patient's real needs based on patient parameters of the existing case studies not having used BIS. Regarding the shortage of BIS in case hospitals and the urgent need to improve anesthesia management, this paper focused on estimating the drug dose without the use of BIS. In this study, a different approach is applied to design an adaptive estimator for computing the infusion propofol dose base on the actual patient status. This study aimed to present the newly developed system for clinical practice applying the propofol dose estimation model.

This paper focuses on regulating the depth of anesthesia (DOA) by estimating the proper dose of propofol as a hypnotic agent without the use of BIS. The proposed ANFIS estimation model under the direct supervision of an anesthesiologist compares the results from control monitoring tool guided (BIS) and the PK-PD model in terms of clinical feasibility and accuracy. This model may allow surgeons or anesthesiologists to validate the model results or contribute to fine-tuning.

In this regard, the models are not dependent on BIS data, especially when the system is affected by multiple related patient parameters. To practically assess the propofol dose estimation model on the patients' depth of anesthesia, a real case study was carried out in Mashhad hospitals in Iran. To deal with the uncertainty of the proposed model, an integrated neural network with the fuzzy inference system was proposed as an ANFIS model. It deals with uncertain data when training and designing a model to estimate the drug dose according to the patient's real needs in the presence of multiple related events. The proposed estimation model aims to optimize anesthesia drug dose to diminish the total drug dose usage, minimize surgical costs, and reduce recovery time.

In this paper, the main objectives are summarized below:

- Appling the estimation model of the depth of anesthesia in the absence of BIS according to the previous work;
- Proposing a propofol dose estimation based on the DOA of real patients without the BIS monitoring;
- Considering uncertainty of patients' parameters in training and designing model by applying ANFIS to estimate propofol dose during the induction and maintenance separately.
- Introducing a multi-parameter decision-making system to anesthesiologist to tackle the challenges in anesthesia administration.

The rest of this paper is organized as follows. The research methodology is presented in section 2. Section 3 provides a comprehensive analysis of the results. The managerial implications are discussed in section 4. Finally, the concluding remarks are recommended for future work in section 5.

2. METHOLOGY

As a case study, an Iranian care public hospital was adopted in this study. As mentioned, this paper focuses on the anesthetic drug dose and the hypnosis control problem. The level of unconsciousness and hypnosis cannot be directly measured by the use of Bispectral Index (BIS), which is commonly used for indirect measurement. The patient's vital signs correlated with hypnosis are derived from the electroencephalogram (EEG) signals [16]. BIS is a dimensionless index that measures the level of unconsciousness with values ranging from 100 to 0 (100 = awake, 0 = no brain electrical activity). In general anesthesia, a BIS signal varies from 40 to 60. The problem in the closed-loop controller is calculation or estimation of the optimal drug dose during anesthesia for maintaining the patient's unconsciousness at an adequate level. A well-designed system must be able to refuse any disturbances (surgical stimulus, blood loss, etc.) during surgery. Generally, a value of 50 is considered as a set point appropriate for the surgical procedure [16]. Since there was limited access to BIS in our case study, DOA was monitored based on previous work [17]. Monitoring of DOA provides the feedback signal port with a laptop connection. Software

was hosted on the laptop to implement the algorithm of closed-loop control. Figure 1 represents a patient diagram in the operating theater for the closed-loop system to estimate the propofol dose per patient. In the pre-anesthesia phase, the patients' initial physiological parameters and information were collected. Then, the types of anesthesia and anesthetic drug as well as the dose of drug were determined by the anesthesiologist and the surgeon.

Data were collected on the patient's preparation before the surgery from the induction phase of anesthesia that begins with the injection of an estimated dose of propofol. An automatic propofol injection pump was connected to the patient based on the depth of anesthesia in such a way that anesthesia was administered by either the required amount of propofol injection into the patient's body or the stopped injection. About 5 to 7 minutes after the induction phase, patients entered the maintenance phase. This process is a closed-loop control system that injects the necessary propofol based on the patient's actual needs using an automatic injection pump connected to the patient. The estimator controls the closed-loop to prevent system failures with an alarm module. Errors in connection or transmission between various devices refer to the low-quality signal of DOA alerted by this module. The DOA was continuously monitored by this program, the suppression rate, and the electromyogram. According to the signal values, the alarm module was activated by the controller to alert the anesthesiologist. Besides, stop mechanism injection for the DOA below a given value (40) and a bolus dose injection for the DOA excessively high value (60) are assigned in the closed-loop system as safety measures to prohibit eventual consciousness. Thus, the neuro-fuzzy estimator model was designed to predict and control the patient's propofol dose based on input data, monitoring the vital signs every minute. The propofol dose estimation model will result in safe anesthesia with the minimum amount of medication required by the patient.

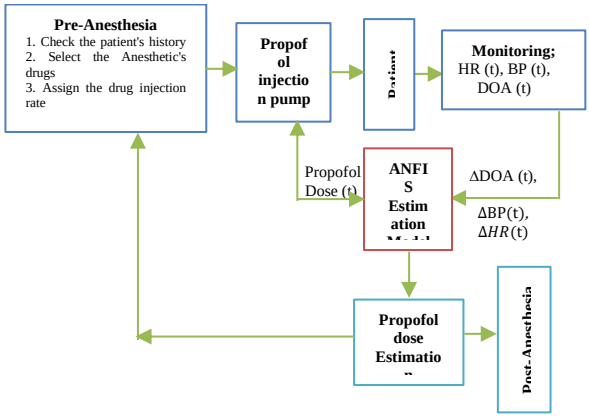


Figure 1. ANFIS model Propofol dose estimater

2. 1. Patient Model The model of the patient explains the relation between drug effect and drug infusion rate through a neuro-fuzzy inference system. On the one hand, the estimative model should be able to meticulously predict the dynamic responses of patients to the rate of infusion. On the other hand, it must not be complicated to avoid time-consuming calculation. In this model, the reaction of the patient undergoing anesthetic to the propofol dose is simulated and estimated based on the patient's needs using artificial intelligence techniques.

In the estimation propofol dose model, the patients' physiological parameters enter the model as inputs. The components of inputs are height, weight, age, gender, heart rate (HR), blood pressure (BP), type of surgery, surgery duration based on surgeon approval, and the depth of anesthesia (DOA). The proposed model structure is shown in Figure 2. The training model begins with the input data given to the system based on the actual data of patients during the operation. The single output of the model is the propofol dose estimated based on the patient's real needs every minute. The paper provides a training model in the induction phase, then models the maintenance phase according to the anesthesia process. The proposed model was finalized based on the propofol dose estimation for 5 to 7 minutes in the induction phase and during surgery in the maintenance phase. To validate the model, the results were compared with the real data and that the classical (PK-PD) model.

2. 2. Patient Data The data were collected from patients of an Iranian hospital whose informed consent was fully obtained before the operation. Data collection was done by observing anesthetized patients during surgery. According to the healthcare unit protocols, ethical authorization was needed to monitor the patients' physiological data. After receiving official approval to observe patients under surgery in the operating room of the study hospital, the patients were tracked in all stages from pre-anesthesia to post-anesthesia. Before surgery, the surgeon consulted with the cases, all of them were

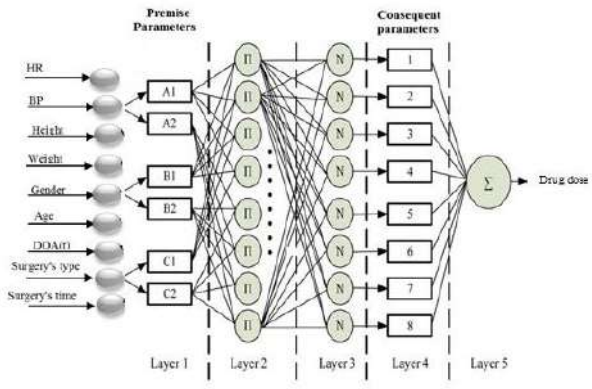


Figure 2. ANFIS propofol estimation model

elective surgery patients. In pre-anesthesia, patients' parameters of age, gender, height, weight, BP, HR, and level of consciousness were recorded. According to the case history and operation type in pre-anesthesia, the surgeon assigned the approximate operation time. Therefore, the anesthesiologist determines the average drug dose through consulting the surgeon.

Concerning the protocol of medical, standard monitors including pulse oximetry, noninvasive blood pressure, and ECG were daily used. In pre-anesthesia, an intravenous catheter was placed in the patient's arm. Depending on the analgesic requirements, the remifentanil was injected (0.25 mcg/kg/min) just before the induction phase. As long as the patient's hemodynamic status became stable, the clinicians maintained the injection rate .

2. 3. Estimation Model by Adaptive Neuro Fuzzy Inference System (ANFIS) ANFIS is far more intricate than fuzzy inference systems, not accessible by all their options. In ANFIS modeling, the data are first examined under fuzzy (clustering). Next, the data are trained using fuzzy inference of Sugeno or Takagi-Sugeno-Kang and then the error is reduced by the fuzzy rules of IF-THEN and functions of membership. The Sugeno type of fuzzy inference system is created from cluster information using a minimum number of rules to well model the behavior of data. The rules are self-divided regarding the fuzzy qualities of each data cluster. To train the ANFIS model, derivative-based algorithms such as backpropagation utilize each node with differentiable functions. Input parameters are Age, Height, Weight, and Gender. The variables are described in Table 1.

A predefined structure is formed where the related parameters determin each membership function (MF). Then, a learning algorithm adjusts these parameters to the N-sample data set. The Takagi-Sugeno Kang fuzzy rules in the fuzzy inference system are described for ANFIS as follows [17]:

TABLE 1. Patient inputs/ output variables

Inputs Variables	
Blood pressure at t	$BP(t)$
The differential Blood pressure at t	$\Delta BP(t)$
Heart rate at t	$HR(t)$
The differential Heart rate at t	$\Delta HR(t)$
The Depth of Anesthesia at t	$DOA(t)$
The differential the Depth of Anesthesia at t	$\Delta DOA(t)$
Drug Dose	D_s
Outputs Variables	
The differential the Depth of Anesthesia at t	$\Delta DOA(t+i)$

$$R_i: if\ x\ is\ A_i,\quad then\ y = f_i(x) = \sum_{j=1}^{n} a_{i,j}x_j + a_{i,o} \tag{1}$$

$A_i = \{A_i^1, A_i^2, \dots, A_i^n\}$, represent the sets of fuzzy and $a_{i,j}, j = 0, 1, \dots, n$ shows the outcom of the parameters. The system output at the *t* stage will be:

$$\hat{y}_t = \frac{\sum_{i=1}^{K} \mu_{A_i}(x_t) f_i(x)}{\sum_{i=1}^{K} \mu_{A_i}(x_t)} \tag{2}$$

$$E_t = \frac{1}{2}(\hat{y}_t - y_t)^2 \tag{3}$$

The specifications of the ANFIS model are illustrated in Table 2. These results are obtained using all 9 inputs, 4 hidden layers, and a single output. The estimation model is trained based on the patients' data.

2. 4. PK-PD Method The classic model defines the relationship between the drug infusion rate and the drug effect by the pharmacokinetics-pharmacodynamics (PK-PD) model.

Pharmacodynamics (PD) explains the relationship between the site concentration effect of the drug and its clinical effect, while pharmacokinetics (PK) describes the drug effect on the body by the infusion rate, drug distribution, and drug elimination. A schematic diagram is represented in Figure 3.

The three-compartmental general model is:

$$\begin{aligned} \dot{C_1}(t) &= -[K_{10} + K_{12} + K_{13}].C_1(t) + K_{21}.C_2(t) + K_{31}.C_3(t) + u(t)/V_1 \\ \dot{C_2}(t) &= K_{12}.C_1(t) - K_{21}.C_2(t) \\ \dot{C_3}(t) &= K_{13}.C_1(t) - K_{31}.C_3(t) \end{aligned} \tag{4}$$

The drug concentration is represented by C_1 [mg/l]. Muscle and fat are the circumferential compartments that model the drug replacement of the blood with fine and weak sporadic tissues of the body. C_2 and C_3 correspond to the slow and fast equilibrating circumferential compartments drug concentrations. The iterative drug

TABLE 2. The ANFIS model specifications

Parameter	**Description**
Number of nodes	316
Number of linear parameters	1040
Number of nonlinear parameters	40
Total number of parameters	1072
Number of training data pairs	98
Number of testing data pairs	6
Number of fuzzy rules	142

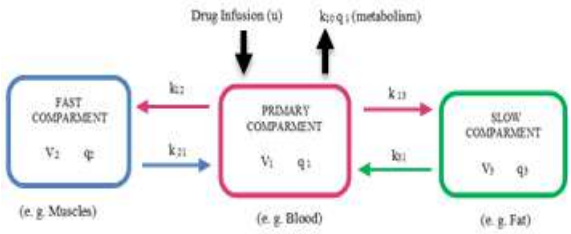


Figure 3. The three-compartmental PK model of patient

transfer from *i*th to *j*th compartment is defined by *kij* (*i*=1, 2, 3), $i \neq j$, and k_{10} implies the drug iteration elimination in the central compartment. The rate of the analgesic drug injection into the central compartment is defined by $u(t)$ [mg/min] [2, 17]. The relevance between drug concentration and the effect of a patient's pharmacology is referred to as PD. The mathematical PD-models are:

$$BIS(t) = E_0 - E_{max} \cdot \frac{C_e(t)^{\gamma}}{C_e(t)^{\gamma} + EC_{50}^{\gamma}} \quad (5)$$

Observing the effect of drug on the patient, the variation of BIS values may depend on the effective C_e, drug concentration by the experimental non-linear Eq. (5) [6-8], named the Hill-Curve in the PD model. In the Hill equation, $Ce(t)$ is the concentration of propofol (μg/ml) in the brain. E_0 expresses the patient's initial state (awake, without medication), which is normally considered to be 100. E_{max} represents drug dose maximum effect, and EC_{50} is the concentration of drug at 50 percent of the maximum effect indicating the medication sensitivity of the patient. The γ shows the curve slope. The set point during anesthesia is 50, with a sufficient level of relaxation provided by values between 40 and 60. Considering the medication effect observed in the patient, the BIS values can relate to the C_e effect and the concentration of the drug (t) based on the experimental nonlinear relationship [9, 17]. To predict the propofol dose using the equations of the PK-PD model, the patients' data were applied and the results were compared with the output of the proposed model.

3. RESULTS AND DISCUSSION

In this study, a single output of the model based on the nine input data in every experiment was applied to design and model the ANFIS estimator, presented in Table 1. The results of estimation model illustrate that the intelligent system can strongly estimate the propofol dose. Therefore, a neuro-fuzzy inference system as well as the ANFIS estimation model was applied in this study. The obtained outcome was compared with that of the PK-PK model and the real data. From the proposed model results shown in Table 3, it can be infered that ANFIS had a better performance on modeling.

TABLE 3. The result of propofol dose estimation by ANFIS model in the induction phase

Induction Time (min)	Real Data	ANFIS-estimated Propofol Dose	ANFIS error
1	1.6	1.6001	0.0001
2	1.6	1.6001	0.0001
3	1.6	1.6	0
4	1.6	1.601	0.001
5	1.6	1.6	0
6	1.6	1.6	0
7	1.6	1.6	0

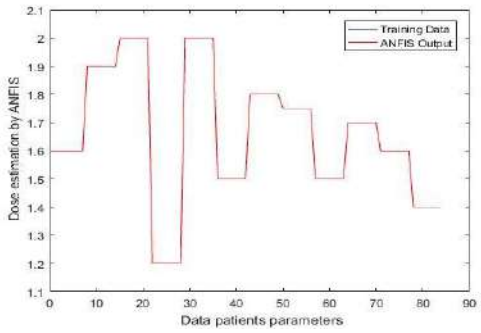


Figure 4. Propofol dose estimation model by ANFIS

Since ANFIS combines a neural network and a fuzzy system to apply experts' knowledge, learning phase implements a more powerful model.

Figure 4 shows the estimation model of propofol dose using a neuro-fuzzy system.

3. 1. ANFIS Estimation Model vs PK-PD Model

Different indicators were used to validate the proposed model, evaluating the error rate and estimating the model accuracy. These evaluation indicators were accuracy, d_α, reliability, MSE, and RSME. The proposed estimation model was compared with the classical PK-PD model and the real values of patients during surgery. Table 4 represents the comparison results of of these indicators between models.

TABLE 4. The comparison of estimation model with real patients' data and PK-PD model in the induction-maintenance phases

Estimation / Method	Propofol dose (induction- maintenance)				
	Accuracy	d_α	Reliability	MSE	RMSE
PK-PD	0.966	0.951	0.95	0.002	0.043
BIS	0.975	0.98	0.99	0.003	0.055
FFNN	0.989	0.999	1	0.001	0.101
Fuzzy	0.953	0.999	1	0.002	0.043
ANFIS	1	1	1	5.3×10^{-6}	0.002

3. 2. Sensitivity Analysis To do sensitivity analysis of the parameters affecting the model, two different senarios were considerd. In each run of the model, by omitting parameter/s of the input data, the final output of the ANFIS model was acquired. The model was also compared with the classical PK-PD method and BIS values to investigate the effect of input on estimating the propofol dose. In first senario, one input parameter was removed and then the model was trained. The outcome was compared with the final output and the estimation results by removing the relevant parameter from the actual values. The results are shown in Figures 5-8, by removing a model input parameter. It is noteworthy that the results were obtained after 20 times implementation of a neuro-fuzzy inference system in MATLAB.

Figure 5 shows the patients' variables and parameters to estimate the required propofol dose in the absence of Age with high accuracy. Therefore, it doesn't have a remarkeable effect on the drug dose

Figure 6 indicates acceptable accuracy of propofol dose estimation without height, so it does not have a considerable effect on the drug dose. Figure 7 shows the gender alone doesn't have a significant effect on the drug dose. Figure 8 shows that there is a significant difference between the actual and the estimated amount of propofol dose in the absence of the weight parameter. As shown by the figures, it is observed that the patient's age has a significant effect on the propofol dose estimation, which

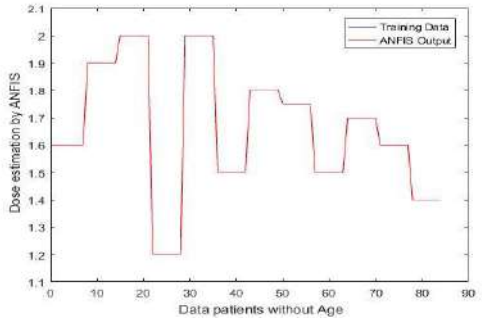


Figure 5. Propofol dose estimation without Age

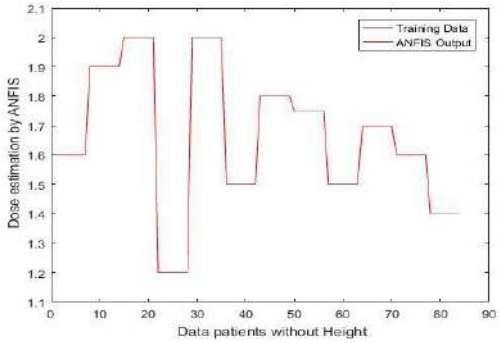


Figure 6. Propofol dose estimation without Height

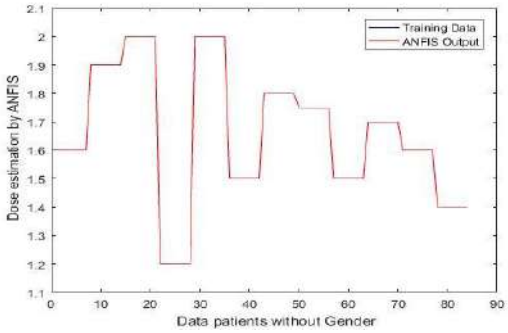


Figure 7. Propofol dose estimation without Gender

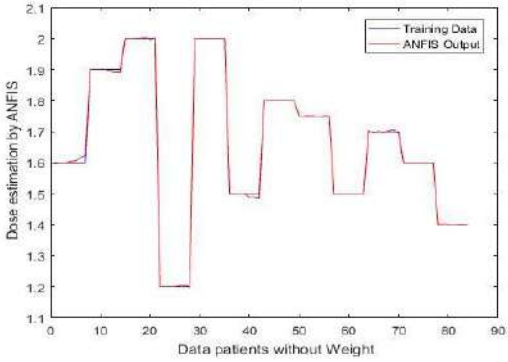


Figure 8. Propofol dose estimation without Weight

can also be affected by the patient's weight compared to other parameters.

Another noteworthy point in the propofol dose estimation is that the patient variables including HR, BP, and DOA in each time should not be omitted in the training model under no circumstances. Otherwise, the estimation results indicate a high error in the absence of vital physiological variables. The results of the sensitivity analysis of parameters affecting propofol dose estimation using ANFIS are summarized in Table 5.

In the second scenario of sensitivity analysis, we examined the effect of patient parameters on the estimation of propofol dose by simultaneous removal of two parameters. The results are shown in Figures 9 to 14.

TABLE 5. The sensitivity analysis of propofol dose estimation parameters by ANFIS

Estimation / Model Parameter	ANFIS				
	Accuracy	d_α	Reliability	MSE	RMSE
Age	1	1	1	3.59×10-4	0.0099
Weight	1	1	0.99	4.52×10-4	0.067
Height	1	1	1	9.29×10-4	0.0031
Gender	1	1	1	3.62×10-4	0.0019
Age	1	1	1	3.59×10-4	0.0099

Figure 9 shows the patients' variables and parameters estimate the required propofol dose in the absence of Age-Gender with accepetable accuracy by slight differences in the drug dose estimation.

Figure 10 represents the patients' variables and parameters estimate the required propofol dose in the absence of age-height with some errors. Therefore, these parameters have a significant effect on the drug dose estimation.

Figure 11 indicates the patients' variables and parameters to estimate the required propofol dose in the absence of height-gender with accepetable accuracy. Therefore, these parameters have a minor effect on the drug dose estimation. Figure 12 shows high differences from the actual data in the absence of age-weight . So, these parameters have a significant effect on the drug dose estimation.

Figure 13 represents the patients' variables and parameters to estimate the required propofol dose in the absence of Weight-Gender with some errors. Therefore, these parameters have a significant effect on the drug dose estimation.

Figure 14 indicate that age and weight are the two parameters with the highest effect on prescribing propofol. The results indicate that age and weight have the highest effect on prescribing propofol, are consistent with those of classical model and BIS values [3, 12, 17].

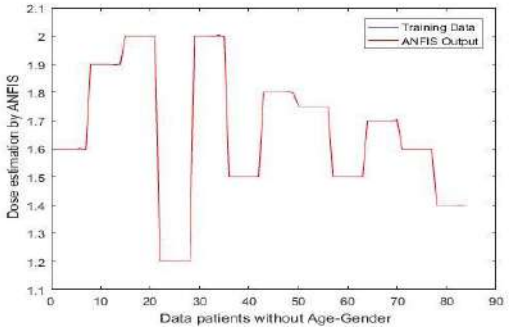


Figure 9. Propofol dose estimation without Age- Gender

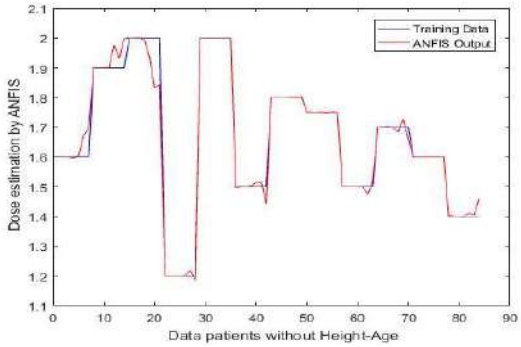


Figure 10. Propofol dose estimation without Age- Height

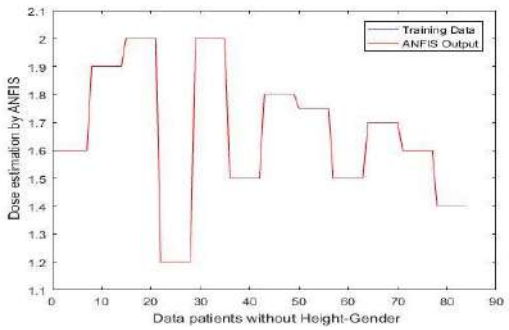


Figure 11. Propofol dose estimation without Height- Gender

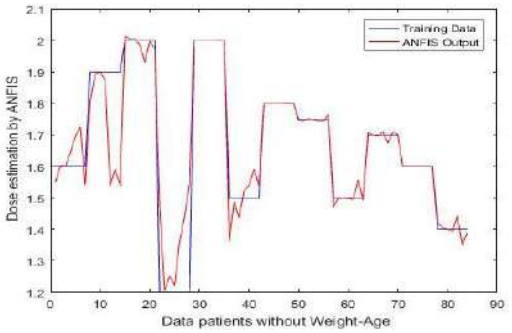


Figure 12. Propofol dose estimation without Age- Weight

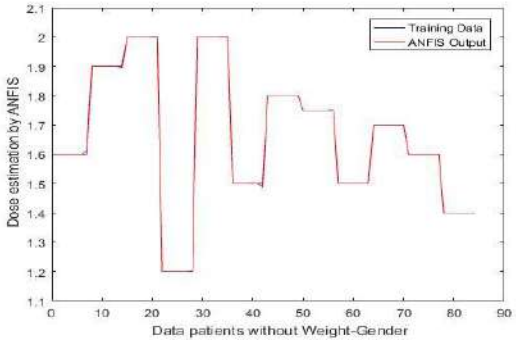


Figure 13. Propofol dose estimation without Weight- Gender

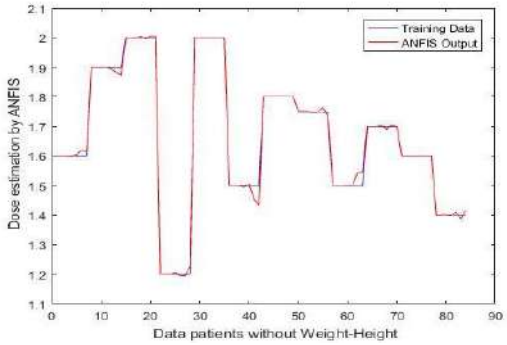


Figure 14. Propofol dose estimation without Weight- Height

4. MANAGERIAL IMPLICATIONS

This paper presents a model for estimating the anesthesia drug dose to cope with possible major anesthesia administration problems, the over/under dosing, propofol wastage in anesthesia, patient's long stay in the hospital, and postoperative complications. This study is motivated by the necessity for safe anesthesia and pain relief in post-anesthesia in hospitals. Since observing data in the operation room needs entry authorization and there is limited number of patients available concerning their historical data. We applied the artificial intelligence techniques to deal with imprecise data for estimation. This neuro-fuzzy training and designing modeling provides enough flexibility regarding input data. The observation of all physiological parameters of the patients at pre-anesthesia, under surgery, and even post-anesthesia stages is of importance for estimating sufficient drug dose. The rationale behind is evident as follows. Each patient reacts differently and his/her parameters may differ from others. In order to satisfy all anesthesiologists' expectations and overcome arising problems from the injected drug, the model was trained based on the patients' parameters to estimate the real dose of propofol needed to be injected. This estimation model aims to compensate for the shortage or limit the surplus of the propofol dose while making reduction in propofol wastage and recovery time. Accordingly, the anesthesiologists prescribe a safe anesthesia delivery during the surgery and post-surgery if the process is managed without the unexpected problems so that every critical situation is predicted directly by the nearest monitoring alarms. This means that the existing anesthesia prediction models have limitation that can just control the parameters which are monitored during surgery.

As can be seen in Figure 4, the estimation model is trained by patients' physiological parameters of age, weight, height, and gender and patients' variables of heart rate, blood pressure, and depth of anesthesia in every minute. The ANFIS model accurately estimates the propofol dose based on the patient's actual needs in comparison with the BIS data. The estimation model is then compared with real patient data and the PK-PD model in the induction and maintenance phases; the results are summarized in Table 5. According to the high uncertainty in the patient's parameters and the effect of drugs on the patient's body, in this study, propofol was considered as the only anesthetic drug neglecting the effect of drugs on each other. The majority of estimation models are designed under this assumption that there is only one drug and they neglect the effect of subsequent drugs.

5. CONCLUSION AND FUTURE WORKS

The results of this work can be compared with those obtained from the PK-PD model and real data of patients. The ANFIS estimation model proposed in this work is data-driven. Thus, the results obtained from propofol dose estimation in the induction and maintenance phases indicate that the presence of data is inalienable for the Artificial Intelligence (AI) estimator to outperform other predictors. ANFIS technique applied in the induction and maintenance phases is more accurate and faster than others. The estimation model with accuracy of 0.999 reduces the total amount of propofol dose. This result is supported by the performance of ANFIS estimator when compared to the PK-PD controller and real data results under the supervision of the surgeon and the anesthesiologist.

Making use of AI estimator, the propofol dose has been thoroughly modeled to regulate the depth of anesthesia during the surgery in this work. The importance of the data-driven estimation model has also been highlighted. Moreover, the role of the drug dose estimator in coping with the noisy signal has been addressed by applying AI techniques. The results of dose estimation show the accuracy of the model in determining the propofol dose in accordance with the real needs of the patient. Furthermore, this dose of anesthetics not only brings safe anesthesia but also minimizes risks of consciousness during surgery or delayed recovery by optimizing the amount of propofol dose. These results will be valuable in clinical trials and this method must be put into practice in hospitals. There are also some suggestions for future studies. The propofol dose optimization is an interesting topic for continuation of this work. The multi-period (pre-anesthesia, anesthesia and post-anesthesia) can be further considered to determine the propofol dose of the estimation model. The costs of drug wastage merits further investigation to be addressed using a cost reduction optimization model.

5. REFERENCES

1. Mu, J., Jiang, T., Xu, X., Yuen, V. and Irwin, M., "Comparison of target-controlled infusion and manual infusion for propofol anaesthesia in children", ***British Journal of Anaesthesia***, Vol. 120, No. 5, (2018), 1049-1055, DOI: 10.1016/j.bja.2017.11.102.
2. Zhang, J. and Huang, C., "Dynamics analysis on a class of delayed neural networks involving inertial terms", ***Advances in Difference Equations***, Vol. 2020, No. 1, (2020), 1-12, DOI: 10.1186/s13662-020-02566-4.
3. Van Den Berg, J., Vereecke, H., Proost, J., Eleveld, D., Wietasch, J., Absalom, A. and Struys, M., "Pharmacokinetic and pharmacodynamic interactions in anaesthesia. A review of

current knowledge and how it can be used to optimize anaesthetic drug administration", ***British Journal of Anaesthesia***, Vol. 118, No. 1, (2017), 44-57, DOI: 10.1093/bja/aew312.

4. Sahinovic, M.M., Struys, M.M. and Absalom, A.R., "Clinical pharmacokinetics and pharmacodynamics of propofol", ***Clinical Pharmacokinetics***, Vol. 57, No. 12, (2018), 1539-1558, DOI: 10.1007/s40262-018-0672-3.
5. van Heusden, K., Soltesz, K., Cooke, E., Brodie, S., West, N., Gorges, M., Ansermino, J.M. and Dumont, G., "Optimizing robust pid control of propofol anesthesia for children; design and clinical evaluation", ***IEEE Transactions on Biomedical Engineering***, (2019), doi: 10.1109/TBME.2019.2898194.
6. Wei, Z.-X., Doctor, F., Liu, Y.-X., Fan, S.-Z. and Shieh, J.-S., "An optimized type-2 self-organizing fuzzy logic controller applied in anesthesia for propofol dosing to regulate bis", ***IEEE Transactions on Fuzzy Systems***, Vol. 28, No. 6, (2020), 1062-1072, doi: 10.1109/TFUZZ.2020.2969384.
7. Jin, W., Zucker, M. and Pralle, A., "Membrane nanodomains homeostasis during propofol anesthesia as function of dosage and temperature", ***Biochimica et Biophysica Acta (BBA)-Biomembranes***, Vol. 1863, No. 2, (2021), 183511, doi: 10.1016/j.bbamem.2020.183511.
8. Hsieh, M.-L., Lu, Y.-T., Lin, C.-C. and Lee, C.-P., "Comparison of the target-controlled infusion and the manual infusion of propofol anesthesia during electroconvulsive therapy: An open-label randomized controlled trial", ***BMC Psychiatry***, Vol. 21, No. 1, (2021), 1-10, doi: 10.1186/s12888-021-03069-6.
9. Lai, H.-C., Lee, M.-S., Lin, K.-T., Huang, Y.-H., Chen, J.-Y., Lin, Y.-T., Hung, K.-C. and Wu, Z.-F., "Propofol-based total intravenous anesthesia is associated with better survival than desflurane anesthesia in robot-assisted radical prostatectomy", ***PloS One***, Vol. 15, No. 3, (2020), e0230290, doi: 10.1371/journal.pone.0230290 .
10. West, N., van Heusden, K., Görges, M., Brodie, S., Rollinson, A., Petersen, C.L., Dumont, G.A., Ansermino, J.M. and Merchant, R.N., "Design and evaluation of a closed-loop anesthesia system with robust control and safety system", ***Anesthesia & Analgesia***, Vol. 127, No. 4, (2018), 883-894, doi: 10.1213/ANE.0000000000002663.
11. Kodama, M., Higuchi, H., Ishii-Maruhama, M., Nakano, M., Honda-Wakasugi, Y., Maeda, S. and Miyawaki, T., "Multi-drug therapy for epilepsy influenced bispectral index after a bolus propofol administration without affecting propofol's pharmacokinetics: A prospective cohort study", ***Scientific Reports***, Vol. 10, No. 1, (2020), 1-9, doi: 10.1038/s41598-020-58460-2.
12. Araújo, A.M., Machado, H., de Pinho, P.G., Soares-da-Silva, P. and Falcão, A., "Population pharmacokinetic-pharmacodynamic modeling for propofol anesthesia guided by the bispectral index (bis)", ***The Journal of Clinical Pharmacology***, Vol. 60, No. 5, (2020), 617-628, doi: 10.1002/jcph.1560.
13. Samadi, F. and Moghadam-Fard, H., "Active suspension system control using adaptive neuro fuzzy (anfis) controller", ***International Journal of Engineering***, Vol. 28, No. 3, (2015), 396-401, doi: 10.5829/idosi.ije.2015.28.03c.08.
14. Lashkenari, M., KhazaiePoul, A., Ghasemi, S. and Ghorbani, M., "Adaptive neuro-fuzzy inference system prediction of zn metal ions adsorption by γ-Fe_2O_3/polyrhodanine nanocomposite in a fixed bed column", ***International Journal of Engineering***, Vol. 31, No. 10, (2018), 1617-1623, doi: 10.5829/ije.2018.31.10a.02.
15. Bahadori-Chinibelagh, S., Fathollahi-Fard, A.M. and Hajiaghaei-Keshteli, M., "Two constructive algorithms to address a multi-depot home healthcare routing problem", ***IETE Journal of Research***, (2019), 1-7, doi: 10.1080/03772063.2019.1642802.
16. Sigl, J.C. and Chamoun, N.G., "An introduction to bispectral analysis for the electroencephalogram", ***Journal of Clinical Monitoring***, Vol. 10, No. 6, (1994), 392-404, doi: 10.1007/BF01618421.
17. Jamali, N., Sadegheih, A., Lotfi, M., Wood, L.C. and Ebadi, M., "Estimating the depth of anesthesia during the induction by a novel adaptive neuro-fuzzy inference system: A case study", ***Neural Processing Letters***, (2020), 1-45, doi: 10.1007/s11063-020-10369-7.

Persian Abstract

چکیده

در این مطالعه، دوز داروی بیهوشی براساس نیاز واقعی بیمار و پارامترهای فیزیولوژیکی آن تخمین زده می‌شود. در بین داروهای بیهوشی داخل وریدی، پروپوفول یکی از پرکاربردترین حین جراحی در مرحله القاء و نگهداری بیهوشی بشمار می‌آید. با استفاده از تکنیک های هوش مصنوعی، دوز پروپوفول بر اساس نیاز واقعی تخمین زده می شود. در این مدلسازی، تخمینگر سیستم استنتاج عصبی- فازی تطبیقی برای محاسبه دوز دارو جهت ارائه بیهوشی ایمن بکار رفته است. مدل بر اساس پارامترهای فیزیولوژیکی بیماران واقعی حین عمل مانند (سن ، وزن ، قد ، جنس) ، فشار خون ، ضربان قلب و سطح هوشیاری بیمار، دوز پروپوفول را در مرحله القا تخمین می‌زند. عملکرد مدل برآورد پیشنهادی با مدل کلاسیک فارماکوکینتیک- فارماکودینامیک و داده‌های بدست آمده از بیماران تحت عمل جراحی مقایسه می‌شود. نتایج نشان می‌دهد که مدل برآورد سیستم استنباط عصبی فازی تطبیقی با دقت ۰.۹۹۹ مقدار کل دوز پروپوفول را به میزان قابل توجهی کاهش می‌دهد. نوآوری مدل پیشنهادی برآورد عمق بیهوشی بیمار در فازهای القاء و نگهداری جداگانه و مستقل از شاخص دوطیفی است. مدل پیشنهادی با مطالعه موردی از بیماران در بیمارستان مشهد- ایران جهت اعتبارسنجی ارائه شده است.

International Journal of Engineering

Journal Homepage: www.ije.ir

Mathematical Modeling and Multi Response Optimization for Improving Machinability of Alloy Steel using RSM, Grey Relational Analysis and Jaya Algorithm

A. Venkata Vishnu*[a], S. Sudhakar Babu[b]

[a] *Research Scholar, Department of Mechanical Engineering, Koneru Lakshmaiah Education Foundation, Guntur, A.P, India*
[b] *Associate Professor, Department of Mechanical Engineering, Koneru Lakshmaiah Education Foundation, Guntur, A.P, India*

P A P E R I N F O

Paper history:
Received 01 June 2021
Received in revised form 10 July 2021
Accepted 02 August 2021

Keywords:
Minimum Quantity Lubrication
Vegetable based Oil Cutting Fluids
Response Surface Methodology
Grey Relational Analysis
JAYA Algorithm

A B S T R A C T

In order to minimise the difficulties associatedwith selecting conventional coolants in any machining, cutting fluids like vegetable based oils can serve as a viable alternative. Vegetable based oils when used in combination with eco-friendly techniques like MQL/NDM can have a major impact in any machining. In the present paper, performance characteristics of surface roughness and tool wear in machining of EN 36 steel alloy under near dry machining conditions/ minimum quantity lubrication using vegetable based oil lubricant is studied. The input parameters like MQL flow rate, speed, feed and depth of cut for 5 levels are used in the CCD approach of Response surface methodology. For improving the machinability of alloy steel and to predit the values a regression equation is designed and developed between the input parameter and the output parameters. A multi-response optimum model for the output responses was also developed using RSM, GRA and JAYA algorithm, It was observed from the experiment results that JAYA algorithm was proved the best multi-response optimization technique when compared to grey relational analysis and RSM.

***doi**: 10.5829/ije.2021.34.09C.13*

1. INTRODUCTION

Machining plays an important role in converting raw material to a desired shape by metal removal in the form of chips. Lot of heat is generated near tool and workpiece [1, 2] interface due to the development of friction between them, where cutting fluids are employed to overcome this effect [3]. Lubrication plays a vital role in cooling tool and work piece and flushing the chips away from the machining area, machining performance of vegetable based coolants compared to conventional coolants have improved thermal conductivity in maintaining the cutting temperature during machining, between the workpiece and tool interface and also reduces the ecological problems associated with the envirоinment [4]. Sustainable manufacturing is one of the recent trends in current industrial economy, as it is eco-friendly, cost effective, waste free, energy efficient etc [5]. Hence an attempt is made in order to reduce the use of lubricant, with the help of one of the sustainable manufacturing technique [6] i.e. utilizing Minimum Quantity Lubrication (MQL). Lubricants accounts around 16 to 20% of the total manufacturing costs [7], Among different techniques available in the reduction of lubricnt flow in machining, researchers are suggesting MQL [8] as a viable alternate; as it reduces the flow of lubricant by spraying the mixture of coolant with air [9]. In MQL machining the compressed air mixed with the coolant [10], where the flow of the air need to maintain in bars and flow of coolant need to maintain below 300ml/h. Several researches have been carried out through MQL technique [11], where as in the present paper an attempt is made in order to study the charecteristics of MQL for different flow rates. The RSM is a statistical and mathematical tool used to develop, optimize and improve a process [12]. RSM composed of design with an aim of determining the optimum functioning of an industrial efficiency, considering least

*Corresponding Author Institutional Email: *venkat666vishnu@gmail.com* (A. Venkata Vishnu)

Please cite this article as: A. Venkata Vishnu, S. Sudhakar Babu, Mathematical Modeling and Multi Response Optimization for Improving Machinability of Alloy Steel using RSM, Grey Relational Analysis and Jaya Algorithm, International Journal of Engineering, Transactions C: Aspects, Vol. 34, No. 09, (2021) 1257-1266

experimental effort [13]. The inputs are known as factors or variables and the outputs known as response that generates by the system [14]. RSM comprises of developing experimental designs, processing of regression model and optimization [15, 16]. In the present paper the RSM methodology is used to develop, optimize and improve the process to minimize the surface roughness and tool wear for the selected variables. A multi response optimization using the GRA, advanced and evolutionary technique Jaya algorithm is developed to check the performance characteristics of the objective function.

2. MATERIAL AND METHODS

The experimentation is carried out through central composite design (CCD) of RSM, In CCD, a design comprises of k factors where distance from axial point to the design center is $\alpha = 2^{k/4}$ [17]. Four independent variables namely MQL flow rate, speed, feed and depth of cut were used for experimentation; hence, based on the input factor k, the value of α is to be considered as 2. The coded input variables with 5 levels are tabulated in Table 1 the output responses selected are tool wear and surface roughness. The experimental design is generated with the help of Minitab 19 software and the sequence of experiments for turning operations is tabulated in Table 2, a total of 31 experiments were performed as per the standard order design sequence and the corresponding results surface roughness and cutting temperature is measured accordingly.

The experiments were performed on high speed CNC machine of LOKESH TL20 Max model CNC Machine shown in Figure 1, The MQL setup was developed using five different "Spray gun" maintaining the flow rates of 50, 100, 150, 200 and 250ml/h with an air compressor maintaining a constant air pressure of 2 bars [9] (layout shown in Figure 3). The cutting tools selected for turning are TNMG Uncoated carbide tool. EN 36 Alloy steel is used as work piece material with carbon content of 0.16%. En 36 is the most widely used Alloy Steel as it has wide applications in manufacturing of gears, shafts, pinions, camshafts and gudgeon pins etc. The dimensions of the work piece, selected for the experiment is 32mm diameter X 150mm length shown in Figure 2. The coolant used is vegetable oil based cutting fluid, which is processed by mixing sunflower oil with triethanol amine and oleic acid, maintained in the ratio of 2:1:2 respectively. The mixture of 40ml of sunflower oil, 40 ml of oleic acid and 20 ml of triethanol amine was taken and stirred thoroughly using a mechanical stirrer, the homogeneous mixture prepared is dissolved in water at a ratio of 1:20.

TABLE 1. Input Variables and Their Levels

Factors	Levels / Units	Lower (-2)	Low (-1)	Centre (0)	High (+1)	Higher (+2)
Mql- Flow Rate (A)	ML/HR	50	100	150	200	250
Speed (B)	RPM	700	900	1100	1300	1500
Feed (C)	MM/REV	0.1	0.3	0.5	0.7	0.9
Depth Of Cut (D)	MM	0.5	1.5	2.5	3.5	4.5

TABLE 2. Central Composite design Experimentation with Surface roughness and Tool wear values

MQL flow rate	SPEED	FEED	DOC	Ra	TW
150	1100	0.5	2.5	2.21	0.44
200	1300	0.3	1.5	1.76	0.31
200	900	0.7	3.5	2.22	0.57
100	1300	0.3	3.5	2.54	0.48
200	900	0.7	1.5	2.12	0.59
100	1300	0.3	1.5	2.51	0.37
200	1300	0.3	3.5	1.96	0.41
200	900	0.3	3.5	2.23	0.42
150	1100	0.9	2.5	2.58	0.61
50	1100	0.5	2.5	2.67	0.55
200	900	0.3	1.5	2.19	0.39
150	1100	0.5	2.5	2.24	0.45
150	1100	0.5	2.5	2.23	0.44
150	700	0.5	2.5	2.31	0.41
150	1100	0.5	4.5	2.61	0.44
200	1300	0.7	1.5	2.03	0.51
150	1100	0.5	2.5	2.2	0.45
100	1300	0.7	3.5	2.45	0.56
150	1100	0.5	2.5	2.29	0.49
250	1100	0.5	2.5	1.84	0.45
150	1100	0.5	2.5	2.21	0.43
100	900	0.3	1.5	2.56	0.4
100	900	0.7	1.5	2.51	0.57
100	1300	0.7	1.5	2.47	0.61
150	1100	0.5	2.5	2.25	0.43
150	1100	0.5	0.5	2.34	0.39
100	900	0.7	3.5	2.51	0.59
200	1300	0.7	3.5	2.03	0.63
150	1100	0.1	2.5	2.21	0.21
100	900	0.3	3.5	2.51	0.46
150	1500	0.5	2.5	2.2	0.51

Figure 1. CNC machine

Figure 2. EN36Alloy steel

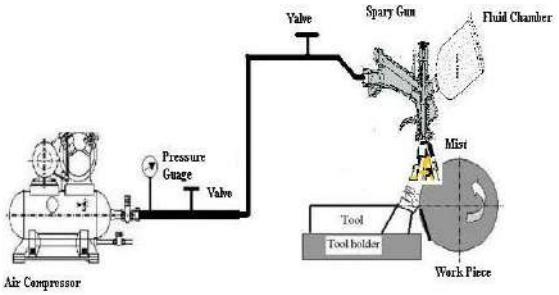


Figure 3. Layout of MQL setup

The Surface roughness (Ra) is measured using MITUTOYO surface roughness tester shown in Figure 4, the results of corresponding Ra values are tabulated in Table 2. Tool flank wear is measured directly using tool makers microscope as shown in Figure 5 at a 100X magnification.

Figure 4. Surface Roughnes tester

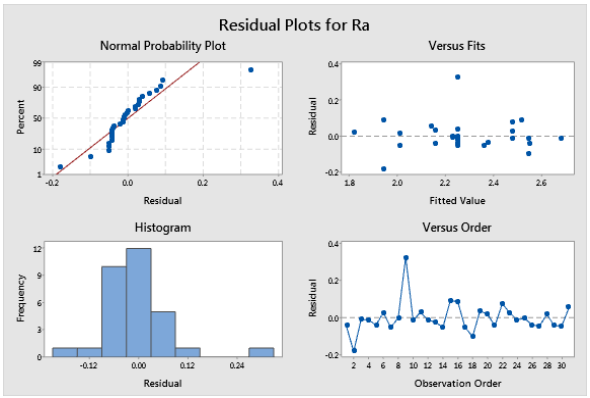


Figure 5. Residual plots for surface roughness regression model

3. RESULTS AND DISCUSSIONS

The machinability affects of Steel alloy using vegetable based cutting fluid considering the input variables is investigated through RSM approach. Table 2 represents the results of measured surface roughness (Ra) and tool wear (Tw) as per the standard order of sequence. In RSM, the experimental design and regression equation helps in retrieving the response for selected independent input variables [14-20] using the following equation:

$$X=b0 + b1Y1 +b2Y2 +b3Y3+\ldots. + bnYn+e \quad (1)$$

where, X is output response, Y1,Y2….. are input factors and its corresponding interactions, and b1, b2….. are the quadratic model associated with regression of RSM.

3. 1. Effect of Input Factors on Surface Roughness

Based on the experimental design the Ra measured, in Table 2, the quadratic equation developed by calculating coefficient of regression for surface roughness is given in Equation (2). The ANOVA is performed, to define the significance of the input variable towards output response and to check the model adequacy, are tabulated in Table 3, model F- calculated value is 14.37 which indicates model is significant. The values of P < 0.0500 imply model terms to be significant. In the present work MQL Flow rate, Speed, Depth of cut, DOC * DOC, MQL flow rate * Speed are said to be significant. A value generated > 0.1 indicates the model is not significant. The lack of fit is 0.4 which indicates it is not significant, as lack of fit with Non-significant is good –as it is needed that the model is to be fit [17]. Model showed a correlation coefficient (R2) of 92.63 % value suggesting a satisfactory representation of model. Furthermore the insignificant model terms are eliminated using backward elimination approach in order to fit the full model, hence the regression equation considering second order terms is given by Equation (3).

$$Ra=3.201- 0.00048\ MQL\ flow\ rate + 0.000394\ SPEED - 1.490\ FEED - 0.310\ DOC- 0.000002\ MQL\ flow\ rate*MQL\ flow\ rate - 0.000000\ SPEED*SPEED + 0.774\ FEED*FEED + 0.0510\ DOC*DOC - 0.000005\ MQL\ flow\ rate*SPEED + 0.00275\ MQL\ flow\ rate*FEED+ 0.000475\ MQL\ flow\ rate*DOC + 0.000531\ SPEED*FEED + 0.000037\ SPEED*DOC - 0.044\ FEED*DOC \quad (2)$$

TABLE 3. ANOVA table of RSM for Surface Roughness

Source	DoF	Adj SS	Adj MS	F-Value	P-Value
Model	14	1.425	0.102	14.37	0
Linear	4	1.248	0.312	44.04	0
MQL flow rate	1	1.118	1.118	157.8	0
SPEED	1	0.073	0.073	10.25	0.006

FEED	1	0.028	0.028	3.95	0.064
DOC	1	0.029	0.029	4.15	0.039
Square	4	0.101	0.025	3.55	0.029
MQL flow rate* MQL flow rate	1	0.000	0.000	0.07	0.8
SPEED*SPEED	1	0.000	0.000	0.07	0.8
FEED*FEED	1	0.027	0.027	3.87	0.067
DOC*DOC	1	0.074	0.074	10.48	0.005
2-Way Interaction	6	0.077	0.013	1.8	0.162
MQL flow rate*SPEED	1	0.046	0.046	6.53	0.021
MQL flow rate*FEED	1	0.012	0.012	1.71	0.21
MQL flow rate*DOC	1	0.009	0.009	1.27	0.276
SPEED*FEED	1	0.007	0.007	1.02	0.328
SPEED*DOC	1	0.001	0.001	0.13	0.726
FEED*DOC	1	0.001	0.001	0.17	0.683
Error	16	0.113	0.007		
Lack-of-Fit	10	0.108	0.011	11.24	0.4
Pure Error	6	0.006	0.001		
Total	30	1.539			

To check the acceptability of reduced model, ANOVA is performed again, but considering the significant terms and tabulated in Table 4. It is observed that the F value shows considerable improvement of 33.08 compared to 14.37 from Table 3. The model displayed at a Confidence level (R^2) of 86.87 %. To validate the regression Equation (3) the input parameters other than the selected values are considered to predict the equation as shown in Table 5. A conformation test is also carried out based on the selected values and the percentage of error is calculated by using the below Equation (4), hence the percentage of error found to be within the range of acceptance i.e. -5.31 to 5.29.

$$Ra=2.534 + 0.00160 \text{ MQL flow rate} + 0.000531 \text{ SPEED} - 0.2097 \text{ DOC} + 0.0489 \text{ DOC*DOC} - 0.000005 \text{ MQL flow rate*SPEED} \quad (3)$$

$$\text{Percentage of error} = \frac{\text{Actual value} - \text{Predicted value}}{\text{Predicted value}} \times 100 \quad (4)$$

To check the adequacy of model, Residual plots are developed for the surface model of Ra shown in Figure 5. The Probability plot of residual values remains on a line, which indicates the experimental values meet the confidence intervals and the guidelines of sample size. In fitted verses residual plot, the residual values are distributed randomly with constant variance and the points are observed on both sides of zero line. In the order verses residual plot the values fall about the center line randomly. From Figure 7, it is evident that the residuals are not independent and thus correlated [17].

TABLE 4. ANOVA table of RSM for modified Surface Roughness

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Model	5	1.333	0.2675	33.08	0
Linear	3	1.222	0.4066	50.31	0
MQL flow rate	1	1.110	1.1180	138.32	0
SPEED	1	0.076	0.0726	8.98	0.006
DOC	1	0.0294	0.0294	3.64	0.048
Square	1	0.0708	0.0708	8.72	0.007
DOC*DOC	1	0.0708	0.0704	8.72	0.007
2-Way Interaction	1	0.0463	0.0463	5.72	0.025
MQL flow rate*SPEED	1	0.0463	0.0462	5.72	0.025
Error	25	0.2027	0.0080		
Lack-of-Fit	19	0.1963	0.0103	10.8	0.4
Pure Error	6	0.0054	0.0006		
Total	30	1.5388			

TABLE 5. Surface Roughness- Validation experiments

	A	B	C	D	Predicted Values	Actual Values	% Error
Exp 1	60	750	0.3	0.4	2.73	2.68	-1.73
Exp 2	120	950	0.6	0.8	2.52	2.39	-5.31
Exp 3	180	1150	0.9	1.2	2.22	2.17	-2.09
Exp 4	240	1350	1.2	1.6	1.8	1.9	5.29

As surface roughness is an output response, which is required to be minimized in any machining operation. Figures 6 and 7 show the 3D response surface and counter plots with interaction effects of process parameters and their effects on the response value; from the surface plots bright spots indicates the effect of surface roughness (Ra) in connection with input parameters. Hence, all the interactions between the variables, especially the effect caused with respect to MQL flow rate to speed, feed and depth of cut are to be more systematic when compared. From counter plots to obtain minimum surface roughness the suggested MQL flow rate lies above 200 ml/h, speed above 1200rpm, feed below 0.4mm/rev and depth of cut between 1 to 3.5mm.

3. 2. Effect of Input Factors on Tool Wear Based on the experimental design the tool wear is measured and the quadratic equation developed by calculating

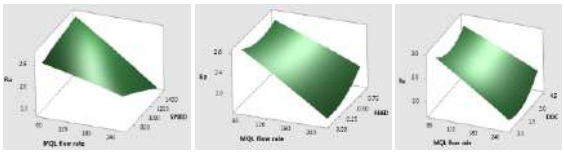


Figure 6. (a) 3-D Surface plot of surface roughness on, MQL-flow rate vs speed (left),MQL-flow rate vs feed (center) and MQL-flow rate vs depth of cut (right)

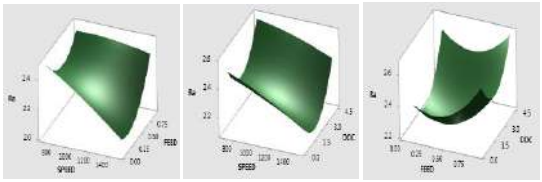


Figure 6. (b) 3-D Surface plot of surface roughness on, speed vs feed (left), speed vs depth of cut (center) and feed vs depth of cut (right)

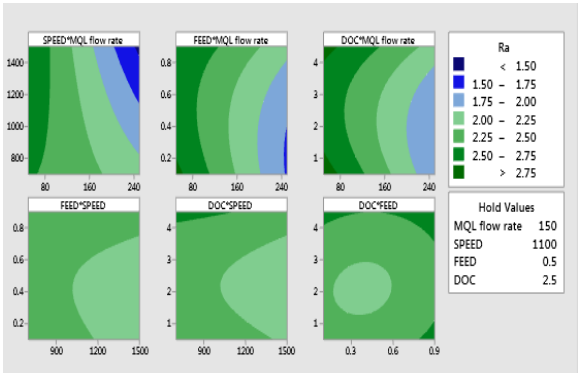


Figure 7. Counter plots for surface roughness with interaction of process parameters

coefficient of second order regression at confidence level (R2) of 84.52% with significant terms is given by Equation (5) [17]. The plots of residual the developed model of tool wear is plotted and shown in Figure 9. As it is observed, the results are shown shown in good arrangement.

$$TW=0.4091 - 0.002748\ MQL\ flow\ rate + 0.4563\ FEED + 0.01958\ DOC + 0.000008\ MQL\ flow\ rate*MQL\ flow\ rate \quad (5)$$

A conformation test is performed to validate the tool wear regression Equation (5) as shown in Table no. 6. The percentage of error found to be within the range of acceptance i.e. -13 to +14. Figures 9 and 10 show the 3D response surface plots and counter plots with the interaction effect of process parameter for tool wear, all the interactions between the variables, especially the effect caused with respect to MQL flow rate to speed, feed and depth of cut are to be more systematic compared with other effects. From counter effects it can be observed that to get minimum tool wear the MQL flow rate above 100 ml/h, speed in between 900 to 1500 rpm, feed less than 0.3 mm/rev and depth of cut less than 2mm is suggestable.

3. 3. Formulation of Multi Objective Function The optimization of two responses namely tool wear and surface roughness in machining of alloy steel under MQL conditions considering the process parameters is studied

TABLE 6. Validation experiments for tool wear

	A	B	C	D	Predicted Values	Actual Values	% Error
Exp 1	60	750	0.3	0.4	0.42	0.48	14.9
Exp 2	120	950	0.6	0.8	0.48	0.42	-13.22
Exp 3	180	1150	0.9	1.2	0.61	0.57	-6.22
Exp 4	240	1350	1.2	1.6	0.79	0.82	3.89

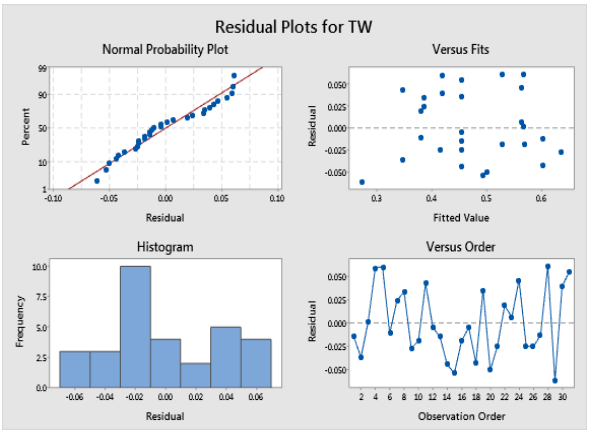


Figure 8. Residual plots for Tool Wear regression model

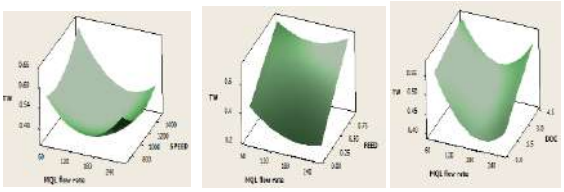


Figure 9. (a) 3-D Surface plot of Tool wear on MQL-flow rate vs speed (left), MQL-flow rate vs feed (center) and MQL-flow rate vs depth of cut (right)

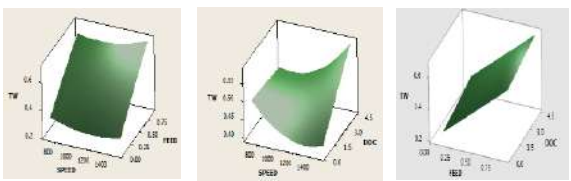


Figure 9. (b) 3-D Surface plot of surface roughness on speed vs feed (left), speed vs depth of cut (center) and feed vs depth of cut (right)

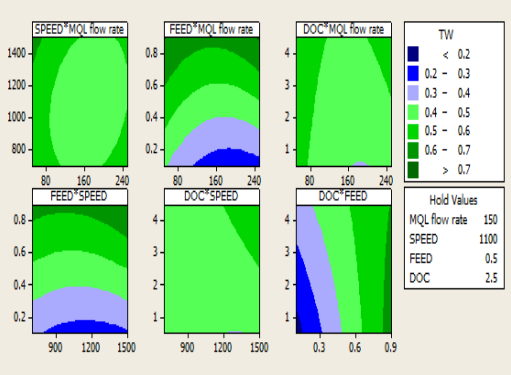


Figure 10. Counter plots of tool wear with interaction of process parameters

and simultaneously a multi objective is formulated for minimization of the output responses. To study the multi objective function a combined objective function is generated to convert multi objective to a single-objective mathematical optimization function, which is given by Equation (6).

$$\text{Min COF} = W1 * (Ra/ Ra\ min) + W2 * (Tw/ Tw\ min) \quad (6)$$

where, W1 and W2 indicates the weights granted to the responses and assigned equal weights of 0.5. After individual optimization the minimum surface roughness adopted to be 1.76μm and tool wear 0.21 mm from Table 2. The normalized multi-objective function to a single objective function is obtained from Equations (3) and (5) is given as Equation (7). The input variables selected are minimum and maximum values of MQL flow rate, speed, feed and depth of cut.

50≤MQL flow rate≤250
700≤Speeed≤1500
0.1≤feed≤0.9
0.5≤depth of cut≤4.5

$$\text{COF}=1.693-0.005\ \text{MQL flow rate} + 0.00014\ \text{SPEED} + 1.085\ \text{FEED} - 0.0132\ \text{DEPTH OF CUT} + 0.0000190\ \text{MQL flow rate}^*\ \text{MQL flow rate} - 0.00000142\ \text{MQL flow rate}^*\ \text{SPEED} + 0.013\ \text{DEPTH OF CUT}^*\ \text{DEPTH OF CUT} \quad (7)$$

JAYA Algorithm

JAYA [18] (Victory in Sanskrit), is an evolutionary optimization technique formulated for solving constrained and unconstrained optimization problem developed by Rao [19]. The Algorithm is based on the concept of shifting towards best solution by avoiding worst solution. For optimization unlike other algorithms, Jaya algorithm requires only basic idea on terms like design variables, objective function, population size and no. of iterations. Figure 11 shows the flow chart of JAYA algorithm, in the present work the objective function is considered to be Equation (7) for minimization of surface roughness and tool wear. The iteration i, with m number of input factors j=1, 2, 3, 4 (MQL flow rate, speed, feed and depth of cut) for population k=1, 2, 3… is considered to modify the best and worst solutions using the Equation (8) [20].

$$A'_{j,k,i} = A_{j,k,i} + r1,j,i\ ((Aj,b,i) - |Aj,k,i|) - r2,j,i\ ((Aj,w,i) - |Aj,k,i|) \quad (8)$$

where Aj,b,i and Aj,w,i is the input variable j for the corresponding best and worst function at i^{th} iteration. A' j,k,i is the modified solution of A j, k, i and r1,j,i and r2,j,i are two random numbers [21]. The Random numbers within the range of input variables are considered and corresponding combined objective function (COF – Equation (8)) is calculated and tabulated in Table 7 for a population size of 6. In the present work the objective is to minimize COF, hence the first row (smallest value) is marked as best solution and fifth row (highest value) is marked as worst. Latter the solution is modified using Equation (8) considering all the variables for each row and column and tabulated in Table 8. On comparing initialize and modified solution row wise the best solution is opted for the 1st iterations, here the modified solution is better than iteration solution hence modified values are considered to be best solutions. As the iterations are continued till final objective has no changes and the last value is considered to be optimum solutions. Using matlab, considering the JAYA algorithm code, a program is run to simulate at various plans (i.e. A to F).

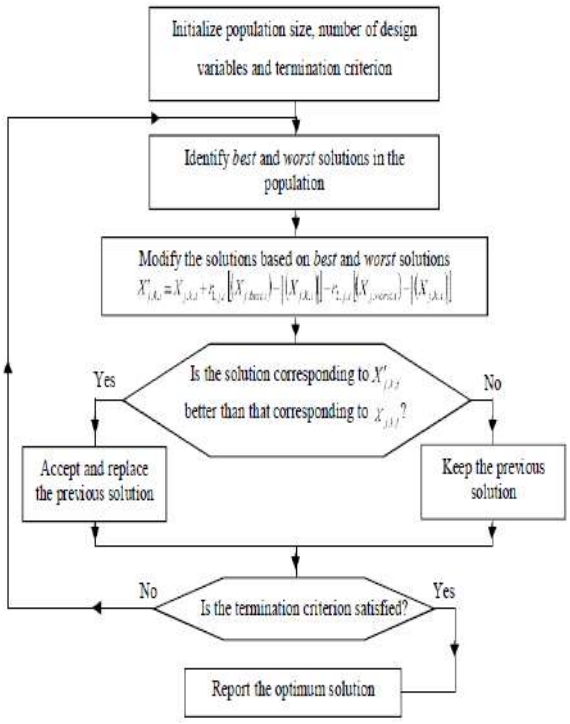


Figure 11. Flow chart of JAYA algorithm

TABLE 7. Initialization of solution

MQL Flow Rate (A)	Speed (B)	Feed (C)	Depth of Cut (D)	COF	
120	1000	0.25	1.2	1.551	Best
170	870	0.38	1.3	1.721	
130	1300	0.48	1.9	1.849	
160	1400	0.52	2.6	1.875	
230	1125	0.62	3.4	2.116	worst
225	1325	0.33	4	1.805	

TABLE 8. Modified and Best solution

A	B	C	D	COF
87	962.5	0.14	0.54	1.4977
132	845.5	0.24	0.47	1.5767
96	1232.5	0.33	1.01	1.7463
123	1322.5	0.36	1.64	1.7257
245	1280	0.4	2.9	1.8474
181.5	1255	0.19	2.9	1.5419

TABLE 9. Performance of machining parameters using JAYA algorithm

Plan	A	B	C	D	COF	Ra	Tw
A	245	1280	0.4	2.9	1.847	1.84	0.46
B	245	1100	0.22	2	1.644	1.94	0.36
C	245	1280	0.22	2	1.607	1.81	0.36
D	245	1460	0.22	1.1	1.545	1.74	0.34
E	245	1460	0.1	2.9	1.485	1.72	0.32
F	245	1460	0.1	1.1	1.415	1.74	0.28

For Plan A, Population size=6, No. of iterations= 1
For Plan B, Population size=10, No. of iterations= 5
For Plan C, Population size=15, No. of iterations= 10
For Plan D, Population size=20, No. of iterations= 20
For Plan E, Population size=15, No. of iterations= 25
For Plan F, Population size=15, No. of iterations= 25

From Table 9, the individual responses of surface roughness and tool wear is also calculated from regression Equations (3) and (5) for various plans and it is observed that the COF, Ra and Tw for the last plan F is very minimum (as our objective is to minimize) and it is considered to be optimum solution.

3. 4. Optimization by Grey Relational Analysis (GRA) From Figure 12 the procedure for generating optimum parameters for multi response optimization using grey relational analysis [22] is shown, the corresponding grey relational coefficient of Ra and Tw values are shown in the Table 10 to minimize the response [23].

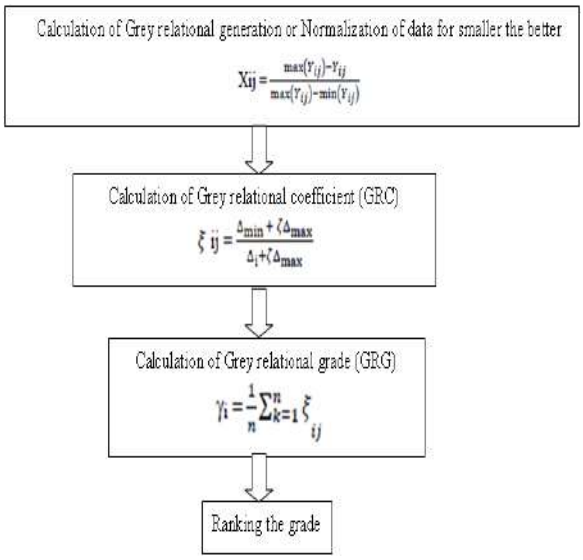


Figure 12. Flow chart of Grey Relational Analysis

TABLE 10. Grey Relational Analysis for Ra and T_w

Exp.No	Normalized values		Grey relational coefficient		GRG	GRG S/N ratio	Rank
	Ra	Tw	Ra	Tw			
1	0.45	0.51	0.48	0.5	0.49	6.2	9
2	0.76	1	0.68	1	0.84	1.53	1
3	0.14	0.49	0.37	0.5	0.43	7.27	22
4	0.36	0.14	0.44	0.37	0.4	7.89	25
5	0.1	0.6	0.36	0.56	0.46	6.8	19
6	0.62	0.18	0.57	0.38	0.47	6.51	17
7	0.52	0.78	0.51	0.69	0.6	4.39	4
8	0.5	0.48	0.5	0.49	0.5	6.09	7
9	0.05	0.1	0.34	0.36	0.35	9.1	31
10	0.19	0	0.38	0.33	0.36	8.93	30
11	0.57	0.53	0.54	0.51	0.53	5.58	5
12	0.43	0.47	0.47	0.49	0.48	6.44	16
13	0.45	0.48	0.48	0.49	0.48	6.29	13
14	0.52	0.4	0.51	0.45	0.48	6.33	14
15	0.45	0.07	0.48	0.35	0.41	7.68	24
16	0.29	0.7	0.41	0.63	0.52	5.69	6
17	0.43	0.52	0.47	0.51	0.49	6.24	11
18	0.17	0.24	0.38	0.4	0.39	8.26	26
19	0.33	0.42	0.43	0.46	0.45	7.03	20
20	0.43	0.91	0.47	0.85	0.66	3.63	3
21	0.48	0.51	0.49	0.5	0.5	6.1	8
22	0.55	0.12	0.53	0.36	0.44	7.06	21
23	0.14	0.18	0.37	0.38	0.37	8.57	27

24	0.05	0.22	0.34	0.39	0.37	8.7	28
25	0.48	0.46	0.49	0.48	0.48	6.29	12
26	0.57	0.36	0.54	0.44	0.49	6.21	10
27	0.1	0.18	0.36	0.38	0.37	8.71	29
28	0	0.7	0.33	0.63	0.48	6.37	15
29	1	0.51	1	0.5	0.75	2.48	2
30	0.4	0.18	0.46	0.38	0.42	7.6	23
31	0.29	0.52	0.41	0.51	0.46	6.74	18

The highest value of GRG obtained through grey relational coefficient, considered as the stronger relational degree and the ranking is obtained accordingly, it is observed that experiment no.2 obtained 1st rank with highest GRG. The optimum level of input factors is determined using results of GRG S/N ratio. Table 11 shows the optimum levels for machining at MQL flow rate 250ml/h, Speed 1300 rpm, Feed 0.1 mm/rev and depth of cut 0.5 mm, where feed ranked with the highest delta value followed by MQL Flow rate, Depth of cut and Speed. The predicted response is calculated as per the Equation (9) which is in good arrangement when compared with the confirmation test results tabulated in Table 12.

$$\text{Predicted Response} = A5 + B4 + C1 + D1 - 3 * (Yij) \quad (9)$$

A5, B4, C1 and D1 are the corresponding input parameters of GRG, Yij- Average of GRG.

3. 5. Optimization by RSM The multi response optimization using perturbation curve (shown in Figure 14) of response surface methodology (RSM) is carried out through minitab 19 statistical software and the optimum values are tabulated in Table 13.

3. 6. Comparisons of Confirmation Test Results The multi response optimization is performed in order to improve the performance characteristics using grey relational analysis, response surface methodology and JAYA algorithm. A confirmation test is performed to

TABLE 11. Response Table for GRG S/N

Level	MQL flow rate	Speed	Feed	DOC
1	8.933	6.331	2.483	6.213
2	7.912	7.209	5.83	6.302
3	6.417	6.343	6.556	6.226
4	5.463	6.167	7.546	7.073
5	3.628	6.743	9.105	7.682
Delta	5.305	1.042	6.622	1.469
Rank	2	4	1	3

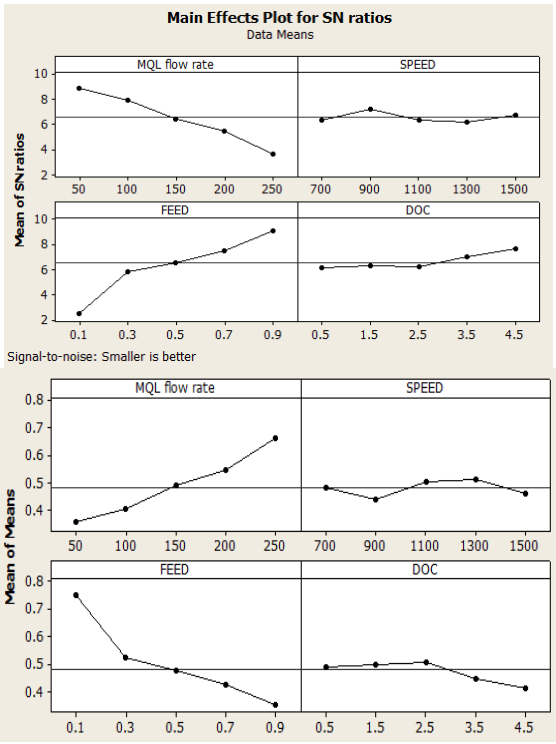


Figure 13. S/N Ratios and Mean plots of GRG

TABLE 12. Confirmation test results of GRA

	Best parameters value out of 31 experiments with GRG are considered to be initial parameters	Optimum parameters	
		Predicted values	Experiment values
Level	A4, B4, C2, D2	A5, B4, C1, D1	A5, B4, C1, D1
Surface Roughness	1.76		1.72
Tool Wear	0.31		0.27
GRG	0.84	0.96	0.92

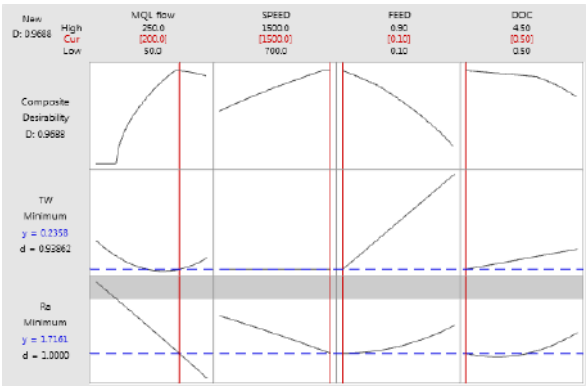


Figure 14. Multi response optimized values with Perturbation curve

TABLE 13. Multi Response optimized values using RSM

Optimum Solution	MQL flow rate	Speed	Feed	DOC
	200	1500	0.1	0.5

validate the model. From Table 14, confirmation test is performed for all the optimum parameters generated through various techniques. The objective is to minimize the dependent variables; it is observed that there was a considerable reduction of surface roughness and tool

TABLE 14. Comparisons of Confirmation test results of JAYA, GRA &RSM

	Initial Parameters	JAYA	GRA	RSM	Change in the results in percentage for the optimum cutting Conditions towards initial parameter settings		
					JAYA	GRA	RSM
MQL Flow rate	200	245	250	200			
Speed	1300	1460	1300	1500			
Feed	0.3	0.1	0.1	0.1			
Depth of cut	1.5	1.1	0.5	0.5			
Ra	1.76	1.67	1.72	1.7	5.11% reduction	2.27% reduction	3.40% reduction
Tw	0.31	0.24	0.26	0.29	22.58% reduction	16.12% reduction	6.45% reduction

wear when compared with initial parameters for all the techniques. The initial parameters selected are the best parameters values out of 31 experiments. The JAYA algorithm shows a reduction percentage of 5.11 % surface roughness and 22.58% toolwear when compared to grey analysis 2.27% Ra and 16.12% Tw, RSM 3.40% Ra and 6.45% Tw.

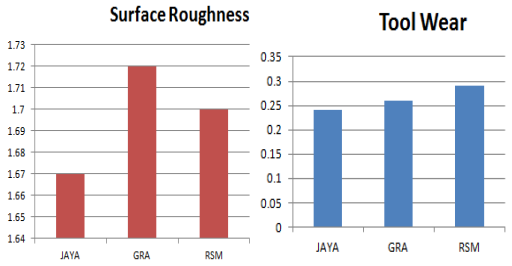


Figure 15. Comparison of confirmation results of Ra and Tw among all the techniques

4. CONCLUSION

The present paper focuses on minimization of surface roughness and tool wear in order to improve the machining of alloy steel under MQL conditions using RSM, GRA and JAYA algorithm techniques. RSM Methodology is implemented and validated successfully in order to study the effect of variables; a quadratic model is developed for surface roughness and tool wear individually and experiments have carried out to confirm the accuracy of the developed model, From the results, it can be concluded that response surface methodology model can predict and develop any output response successfully. From 3D surface and counter plots, there was a considerable impact on the selected independent variables with respect to dependent variables where MQL Flow rate and depth of cut has major impact compared to other variables. Further, the performance characteristics of Surface roughness and Tool wear are identified by multi response optimization using Perturbation curve of RSM Methodology, Grey relational analysis and JAYA algorithm. A confirmation test was performed. At the obtained optimum conditions and compared, the optimum parameters of JAYA algorithm showed better reduction in minimization of surface roughness and tool wear (Figure 15). It is also concluded that machining using MQL at a flow rate of more than 200ml/h gives better result for individual and multi response optimization using any technique.

4. REFERENCES

1. Kryvyi P, Dzyura V, Tymoshenko N, Maruschak P, Nugaras J, Prentkovskis O, "Probability-Statistical Estimation Method of Feed Influence on As-Turned Finish of Steels and Non-Ferrous Metals" ***MDPI-Metals,*** (2018), Vol. 8, No. 11, 965. https://doi.org/10.3390/met8110965.
2. Dzyura, V.; Maruschak, P. "Optimizing the Formation of Hydraulic Cylinder Surfaces, Taking into Account Their Microrelief Topography Analyzed during Different Operations" ***MDPI-Machines,*** (2021), Vol. 9, No. 116. https://doi.org/10.3390/ machines9060116.
3. Bruni, C., Gabrielli, F.F., Simoncini, M., "Effect of the lubrication-cooling technique, insert technology and machine bed material on the work part surface finish and tool wear in finish turning of AISI 420B". ***International Journal of Machine Tools and Manufacture,*** Vol. 46, (2006), 1547-1554. doi.org/10.1016/j.ijmachtools.2005.09.007.
4. M.M.A. Khan, M.A.H. Mithu, N.R. Dhar, "Effects of minimum quantity lubrication on turning AISI 9310 alloy steel using vegetable oil-based cutting fluid", ***Journal of Materials Processing Technology,*** Vol. 209, (2009), 5573-5583

5. Debnath, S., Reddy, M.M., Yi, Q.S., "Environmentally friendly cutting fluids and cooling techniques in machining: a review". ***Journal of Cleaner Production,*** Vol. 83, (2014), 33-47 doi.org/10.1016/j.jclepro.2014.07.071.
6. Chetan, Sudarsan Ghosh, P Venkateswara Rao, "Application of sustainable techniques in metal cutting for enhanced machinability: a review", ***Journal of Cleaner Production,*** (2015) 1-180
7. Anuj Kumar Sharma, Arun Kumar Tiwari, Amit Rai Dixit, "Effects of Minimum Quantity Lubrication (MQL) in machining processes using conventional and nanofluid based cutting fluids: A comprehensive review", ***Journal of Cleaner Production***, Vol. 127 (2016), 1-18. https://doi.org/10.1016/j.jclepro.2016.03.146
8. A. Venkata Vishnu, P. Jamaleswara Kumar, M. Venkata Ramana, "Comparison among Dry, Flooded and MQL Conditions in Machining of EN 353 Steel Alloys-An Experimental Investigation" ***Materials Today: Proceedings***, (2018). doi.org/10.1016/j.matpr.2018.10.296.
9. M. Venkata Ramana, A. Venkata Vishnu, G. Krishna Mohan Rao, D. Hanumantha Rao, "Experimental Investigations, "Optimization of Process Parameters and Mathematical Modeling in Turning of Titanium Alloy Under Different Lubricant Conditions", ***IOSR Journal of Engineering,*** ISSN: 2250-3021 Vol. 2, No. 1, (2012), 086-101.
10. A Venkata Vishnu, P Jamaleswara Kumar "Investigations on high speed machining of EN-353 steel alloy under different machining environments", IOP Conference Series: Materials Science and Engineering, Vol. 330, DOI: 10.1088/1757-899X/330/1/0120 85.
11. Murat Sarikaya, Abdulkadir Gullu, "Multi response optimization of MQL parameters using Taguchi-based GRA in turning of difficult to cut alloy Haynes 25", ***Journal of Cleaner Production***, (2014), 1-11.
12. Manish Gangil, M. K Pradhan, "Modeling and optimization of electrical discharge machining process using RSM: A review", ***Materials Today: Proceedings*** Vol. 4, (2017), 1752-1761. doi.org/10.1016/j.matpr.2017.02.017
13. Neelesh Singh, B. C. Routara, Diptikanta Das, "Study of machining characteristics of Inconel 601in EDM using RSM", ***Materials Today: Proceedings,*** Vol. 5, (2018) 3438-3449. https://doi.org/10.1016/j.matpr.2017.11.590
14. R. Davarnejad, R. Pishdad, S. Sepahvand, "Dye Adsorption on the Blends of Saffron Petals Powder with Activated Carbon: Response Surface Methodology***", International Journal of Engineering Transactions C: Aspects*** Vol. 31, No. 12, (2018) 2001-2008. doi: 10.5829/ije.2018.31.12c.02.
15. M. Safari, V. Tahmasbi, P. Hassanpour, "Statistical Modeling Optimization and Sensitivity Analysis of Tool's Geometrical Parameters on Process Force in Automatic Cortical Bone Drilling Process", ***International Journal of Engineering Transactions B: Applications*** Vol. 34, No. 2, (2021) 528-535. doi: 10.5829/ije.2021.34.02b.26.
16. Tahmasbi, V., Ghoreshi, M., Zolfaghari, M. "Temperature in bone drilling process: Mathematical modeling and Optimization of effective parameters.", ***International Journal of Engineering, Transactions A: Basics***, Vol. 29, No. 7, (2016), 946-953. DOI: 10.5829/idosi.ije.2016.29.07a.09
17. Neelesh Kumar Sahu, Atul B. Andhare, "Multiobjective Optimization for improving Machinability of Ti-6Al-4V using RSM and Advanced algorithms", ***Journal of Computational Design and Engineering,*** (2018), doi: https://doi.org/10.1016/j.jcde.2018.04.004.
18. Rao, R.V. (2016). Jaya: A simple and new optimization algorithm for solving constrained and unconstrained optimization problems, ***International Journal of Industrial Engineering Computations***, Vol. 7, No. 1, 19-34, doi: 10.5267/j.ijiec.2015.8.004.
19. Rao, R.V. Rai, D.P., Ramkumar, J., Balic, J; "A new multi-objective Jaya algorithm for optimization of modern machining processes", ***Advances in Production Engineering & Management***, Vol. 11, No. 4, (2016), 271-286. http://dx.doi.org/10.14743/apem2016.4.226
20. Rao, R.V., Waghmare G.G.: A new optimization algorithm for solving complex constrained design optimization problems, ***Journal Engineering Optimization,*** Vol. 49, No. 1, (2017), 60-83. https://doi.org/10.1080/0305215X.2016.1164855
21. Neeraj Agarwal, M.K. Pradhan, Nitin Shrivastava, "A new multi-response Jaya Algorithm for optimisation of EDM process parameters", ***Materials Today: Proceedings,*** Vol. 5, (2018) 23759-23768. https://doi.org/10.1016/j.matpr.2018.10.167
22. Senthil kumar Ramuvel, Suresh Paramasivam, Study on tool steel machining with ZNC EDM by RSM, GREY and NSGA, ***Materials Research Technology***, (2020). https://doi.org/10.1016/j.jmrt.2020.02.015.
23. P. Sivaiah, D. Chakradhar, Performance improvement of cryogenic turning process during machining of 17-4 PH stainless steel using multi objective optimization techniques, ***Measurement,*** (2018), doi: https://doi.org/10.1016/j.measurement.2018.12.094
24. HMT "Production Technology", Tata McGraw Hill, 2004.
25. Phillip j. Ross "Taguchi Techniques for Quality Engineering", Tata McGraw Hill, Second Edition, 2005.
26. Sidney H.Avner "Introduction to Physical Metallurgy", Tata McGraw Hill 2nd Edition.
27. Minitab Statistical Software Features - Minitab. (2011),"Software for Statistics, Process Improvement, Six Sigma, Quality – Minitab.

Persian Abstract

چکیده

به منظور به حداقل رساندن مشکلات مربوط به انتخاب خنک کننده های معمولی در هر ماشینکاری ، برپایه روغنهای گیاهی می تواند به عنوان یک جایگزین مناسب عمل کند. روغنهای گیاهی در ترکیب با تکنیکهای سازگار با محیط زیست مانند MQL/NDM می توانند تأثیر عمده ای در هر نوع ماشینکاری داشته باشند. در مقاله حاضر ، ویژگی های عملکرد زبری سطح و سایش ابزار در ماشینکاری آلیاژ فولاد EN 36 در شرایط ماشینکاری تقریبا خشک/ حداقل مقدار روغن کاری با استفاده از روان کننده روغن گیاهی مورد مطالعه قرار گرفته است. پارامترهای ورودی مانند سرعت جریان MQL ، سرعت ، تغذیه و عمق برش برای ۵ سطح در رویکرد CCD روش سطح پاسخ استفاده می شود. برای بهبود قابلیت ماشینکاری فولاد آلیاژی و پیش تعیین مقادیر ، یک معادله رگرسیون بین پارامتر ورودی و پارامترهای خروجی طراحی و توسعه یافته است. یک مدل

بهینه چند پاسخ برای پاسخهای خروجی نیز با استفاده از الگوریتم RSM، GRA و JAYA توسعه داده شد ، از نتایج آزمایش مشاهده شد که الگوریتم JAYA بهترین تکنیک بهینه سازی چند پاسخ در مقایسه با تجزیه و تحلیل رابطه خاکستری و RSM است.

International Journal of Engineering

Journal Homepage: www.ije.ir

An Improved Hierarchical Control Structure for Robust Microgrid Operation and Seamless Mode Transfer under Linear and Nonlinear Loads conditions

A. Norozpour Niazi, N. Vasegh*, A. A. Motie Birjandi

Department of Electrical Engineering, Shahid Rajaee Teacher Training University, Tehran, Iran

P A P E R I N F O

Paper history:
Received 25 June 2021
Received in revised form 30 July 2021
Accepted 04 August 2021

Keywords:
Adaptive Lookup Table Control
Hierarchical Control
Micro-Grid
Nonlinear Loads
Proportional Controller
Multi-resonance Controller

A B S T R A C T

This paper proposes the improved hierarchical-based control of microgrid based on proportional and multi-resonance controllers to compensate for harmonic distortion of nonlinear loads. Moreover, the probable transition of MG, especially from grid-connected to unplanned islanding and unintentional MG resources outage were studied. In current and voltage controllers of three-phase VSIs which are located in the inner level, the proportional and multi-resonant controllers are implemented. To attain proper decoupled (P-Q) power-sharing, a selective harmonic type virtual impedance, and a droop-based control are implemented at the primary level. Next, to reach better restoration and subsequently, seamless transition in accidental islanding, unintentional MG-DG's outage, and synchronization process, the advanced three-phase SRF-PLL with in-loop MAF along with a simple adaptive lookup table are implemented in the secondary level of the control. The MATLAB/Simulink simulation results verified that the proposed method improved the performance of control, effectiveness, and robustness in upstream or local grid variation.

***doi**: 10.5829/ije.2021.34.09c.14*

NOMENCLATURE

ω_c	Cut-off frequency	$G_p(s), G_q(s)$	Droop PI controller transfer functions
ω_f	Fundamental frequency	$K_{Pp}, K_{Ip}, K_{Pq}, K_{Iq}$	Droop PI controller parameters
$K_{pc}, K_{r,ch}$	Proportional- resonant coefficients of the current controller	$G_c(s)$	Current controller transfer function
$K_{pv}, K_{r,vh}$	Proportional-resonant coefficients of the voltage controller	$G_v(s)$	Voltage controller transfer function
h	Harmonic component	$G_{PWM}(s)$	PWM inverter transfer function
T_s	Time sampling	$G_{cap}(s)$	Capacitor transfer function
C, L, R	Capacitor, inductor, and resistor of LC filter	$G_L(s)$	Inductance transfer function
τ_{ff}	Feed-forward time constant	$G_{ff}(s)$	Feed-forward transfer function
f	Frequency	$R^+_{v,f}$	Fundamental positive sequence of virtual resistance
P	Active power	$R_{v,h}$	Harmonic negative and positive sequence of virtual resistance
Q	Reactive power	$L^+_{v,f}$	Fundamental positive sequence of virtual inductance
$I^+_{o\alpha,f}, I^+_{o\beta,f}$	Fundamental positive sequences of output current	$K_{PF,syn}, K_{IF,syn}$	PI controller parameters of the frequency synchronization loop
$I_{o\alpha,h}, I_{o\beta,h}$	Harmonic negative and positive sequences of output current	$K_{PE,syn}, K_{IE,syn}$	PI control parameters of the voltage synchronization loop
$V_{v\alpha}, V_{v\beta}$	Virtual voltage in αβ frame	$K_{PF,res}, K_{IF,res}$	PI control parameters of the frequency restoration loop
E	Voltage	$K_{PE,res}, K_{IE,res}$	PI control parameters of the voltage restoration loop

*Corresponding Author Institutional Email: n.vasegh@sru.ac.ir (N. Vasegh)

Please cite this article as: A. Norozpour Niazi, N. Vasegh, A. A. Motie Birjandi, 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, Vol. 34, No. 09, (2021) 2167-2179

MG	Micro-Grid	DER	Distributed Energy Resource
IM	Island Mode	MPC	Model Predictive Control
GCM	Grid-Connected Mode	PLL	Phase Locked-Loop
MSOGI	Multiple Second-Order Generalized Integrator	SRF-PLL	Synchronous Reference Frame Phase Locked-Loop
VIL	Virtual Impedance Loop	VSI	Voltage Source Inverter
SHVI	Selective Harmonic Virtual Impedance	MAF	Moving Average Filter
POI	Point of Interconnection	PI	Proportional-Integral
LPF	Low-Pass Filter	PR	Proportional–Resonant
FLL	Frequency-Locked Loop	QSG	Quadrature Signal Generator
CCS	Central control structure	DCS	Decentral Control Strategy
HCS	Hierarchical control structure	THD	Total Harmonic Distortion

1. INTRODUCTION

In recent years, because of environmental problems, energy crises, and rising concerns about traditional fossil energy shortages, and others, renewable energy has drawn significant attention from researchers. Hence, power systems have undergone a revolution to guarantee sustainable development and resolve power resource challenges, especially in remote areas. Moreover, the other choice to integrate different varieties of energy resources and power electronics interfaced with units is Microgrid. As a dependable electrical system, it can work in both island mode and grid-connected mode to support the main grid, remote areas, and sensitive loads [1]. Hence, variations on power system configuration such as the load variation, MG's unit connection/disconnection, and alteration in MG operation mode in the POI are some of the significant challenges. These variations on power system configuration will make frequency, phase angle, and voltage amplitude mismatches. These mismatches may result in the inrush currents and voltage spikes and consequently, the probable transition will occur [2].

Furthermore, central, decentral, and hierarchical control are the relevant control structures that are proposed to the MG control. The first one is a good choice for the MGs with joint points to have cooperation for their targets. Although in a microgrid with CCS, the operational cost is reduced, but high bandwidth for communication links is needed in MG with DER to make assurance a satisfactory dynamic response of the system. Similarly, preparing high-speed communication links among subsets will cost for the MG owners. As the central control strategy is dependent on fault/delay, consequently probable MG failure may result. This feature represented that it has low reliability [3-4]. Whereas, based on local measurements of DER in decentral control strategy, the control of MG has happened, independently. Therefore, it is made plug-and-play capability for the MG's DER. Likewise, a completely decentral control strategy is not possible to control multiple DGs because of the potent coupling through their operations in MG. Then, the robustness of DCS in MG with the distributed system is small [5-6]. Therefore, the HCS is proposed to enhance efficiency, reliability, control capability, and operation cost. Also, it can use different resources based on different capacities, topologies, and technologies organized in MG [7-9].

As the MG can operate in two modes of operation; grid-connected (*P*-*Q*) and islanded (*V*-*f*) modes, then the operational control mode is important. In grid-connected mode, the MG is responsible to generate the demanded power into the upstream network. But for the remote areas and in a network with failure or major disturbance, to cover sensitive/local loads, the operation mode is changed and switched to unintentional/intentional islanding. So, it should have the capability to control the *V*-*f* of the local region to make uninterruptable support. Consequently, a smooth transition among GCM to IM of MGs is imperative.

The other reason for MG and power system transitions is load type. There are different load types in power system such as linear [10-13], nonlinear [2, 8, 9], balanced [2, 11], unbalanced [3, 14] loads. Two adaptive and smart controllers are planned to regulate the MG frequency and voltage accurately in IM [10]. Moreover, the implemented controllers which are based on modified droop controllers are depending on MPC and H-infinity manners to ensure the smooth transition between IM and GCM. Moradi et al. [11] have studied the loads' variation, their effect on deviations of frequency and voltage, and stability issues. Then, to reach seamless transition and enhance its operation, an adopted technique for operation mode is designed and implemented. Plus, the system robustness is enhanced by gains regulation of the control loops via a fuzzy-based controller. The PR type control was designed in the inner level to track the reference current with zero steady-state error [12]. In addition, to remove the frequency deviation and ensure equal power-sharing, a restoration loop and droop controller are employed. As the power network and microgrids are regularly under linear and balanced load conditions and consequently without any harmonic distortion, to make restoration and synchronization in such system. The only positive sequence of fundamental components were reported in literature [10-14]. A master-slave controller type was proposed by Imran et al. [15]. The master is composed of hybrid battery-diesel resources. As is mentioned heretofore, the communication links with high bandwidth are needed in the central controller which will cost for the owners.

Now, by increasing the presence of the nonlinear loads in the power system, ignoring their effects is inevitable, especially during unintentional islanding and unplanned variation on a grid configuration. Likewise, power quality and after that distortion on voltage and current are the other issues of networks with nonlinear load conditions, which make the implementation of seamless transition more difficult. Because it is imposed on the voltage and power-sharing loops during the operation of a microgrid. Furthermore, under nonlinear load conditions, the SRF control type can include proportional and resonant terms to make harmonic tracking and after that harmonic compensation. Moreover, the PLL is the other key part in MG control, synchronization, and restoration to estimate phase angle, frequency, and voltage amplitude. Then, choosing appropriate PLL is essential to carry out the smooth transition. Indeed, the SRF-PLL is one of the standard PLL and is normally used in three-phase applications, in a power system with a balanced voltage and without any harmonic distortion. But its efficiency and capability in filtering the voltage disturbances in a grid with harmonic distortion and unbalanced voltage are low. Consequently, many advanced SRF-PLLs with high competence to reject the disturbance are introduced which contain different in loop filters and prefilter [16-19]. Here, the synchronous reference frame phase locked-loop with in-loop MAF designed in the last study [8] is applied which has high efficiency in harmonic distortion.

Besides, making decoupled and harmonic power-sharing is the other issue that affects the seamless transition. Therefore, a VIL is needed to achieve decoupled active and reactive power-sharing. Here, SHVI is implemented to have harmonic power-sharing in MG [20-23]. In this study, to extract the fundamental positive sequence and harmonic positive and negative sequence components of the VSIs output current, the MSOGI module is applied, too. The research contributions are:

• Proposing the control scheme based on the hierarchical structure in three-phase VSIs type microgrid and compensation of the harmonic.
• Making decoupled harmonic power-sharing (P-Q) based on fundamental and harmonic components by applying droop control, SHVI, and MSOGI in the level of primary.
• Study the effect of unintentional resources outage on MG.
• Implementing a simple lookup table in adaptive control of voltage and frequency restoration loop to restore the system.
• Finally, achieve smoother transition in the variation of system configuration, especially load connections, accidental DG outage, and variation from grid-connected mode to unintentional islanding mode.

The organization of this paper is: the hierarchical control structure of the microgrid is represented in section 2 and the details of the control levels are stated, there. In this study, based on the load features and needs to compensate harmonic distortion, PR current and voltage controllers are implemented in the inner level. Next, the primary level contains droop control and selective harmonic virtual impedance loops. In addition, the secondary level consists of adaptive-based restoration and synchronization loops on voltage and frequency, which is applied by a simple lookup table and synchronous reference frame phase-locked-loop. After that, the achieved results of the simulation are represented in section 3 which proves the utility and correctness of the hierarchical-based structure of the proposed scheme. It is conclued at the end of this study.

2. MICROGRID HIERARCHICAL CONTROL STRUCTURE

As referred to literature [1, 8, 9], according to ANSI/ISA-95 standard, the structure of MGs hierarchical control can be arranged into four levels: the inner level with two control loops which consists of voltage and current controllers, the primary level that can contain virtual impedance and droop control, the secondary level which comprises restoration and synchronization loops, and finally territory level that includes economically optimal operation, power management, and power market. Here, the concentration of the hierarchical-based structure of the proposed scheme is on the inner, primary, and secondary levels.

2. 1. Inner Control Level As is mentioned before, this level contains two loops and controllers: the interior and exterior loops which comprehend the current and voltage controllers inside of them, respectively. Here, it should be mentioned that to get a faster and better response, the speed of the interior loop is higher than the exterior one. Generally, the two-dimensional frames which are used are the stationary frame (αβ) and synchronous reference frame (dq). As one of the traits of the frame of dq is that it eventuates easier filtering and controlling procedure by making the DC control variables, the dq-SRF is used to transform current and voltage waveforms toward a reference frame synchronously with the power network voltage. Therefore, PI controller is the proper candidate for the dq structure. But this type of controller has a problem in tracking non-dc variables and subsequently failing to remove the steady-state error. So, the PR type control based on the stationary frame is generally preferred especially in the case of the grid with nonlinear load, issues of the power quality in the form of harmonic distortion, and consequently the necessity to make harmonic compensation [3]. As it is presented in Figure

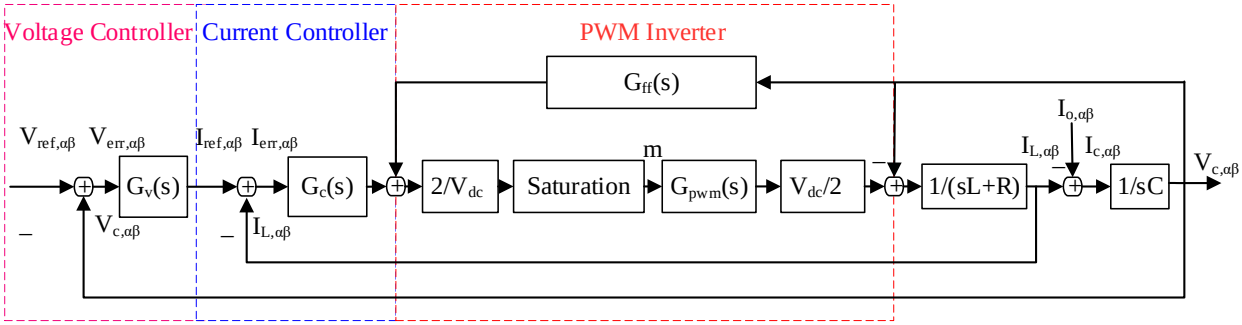


Figure 1. The control scheme of the inner control level of the VSI

1, the inner control level of the VSI has two sections; the LC filters and a microgrid control system with two loops and controllers.

2. 1. 1. Proposed Proportional and Multi-Resonance based Controller in Microgrid and Harmonic Compensation As is mentioned before, the dq and αβ frames are the key-frames that are extensively used in the power system which prepare easier controls and analysis. Since the proportional and integral controller has limitations in tracking non-DC variables and sinusoidal waveforms which resulted in a restriction to remove the steady-state error, the PR controllers are superseded controllers in the case of MG and distributed system in many studies; especially in a power system with nonlinear load condition and subsequently power quality issues [1, 12, 20-25]. As it is represented before, the control scheme of the inner control level of the VSI is presented in Figure 1. Here, in an islanded MG under nonlinear loads situations, the power quality issues should be considered. To implement them, in both current and voltage control loops, the PR-based controllers are proposed to track harmonic components of distorted voltages and currents as follow:

$$G_c(s) = K_{P,c} + \sum_{h=1,5,7,11,\ldots} \frac{K_{r,ch}s}{s^2 + 2\omega_c s + (h\omega_f)^2} \quad (1)$$

$$G_v(s) = K_{P,v} + \sum_{h=1,5,7,11,\ldots} \frac{K_{r,ch}s}{s^2 + 2\omega_c s + (h\omega_f)^2} \quad (2)$$

$$G_{PWM}(s) = \frac{1}{1+1.5Ts} \quad (3)$$

$$G_{cap}(s) = \frac{1}{Cs} \quad (4)$$

$$G_L(s) = \frac{1}{R+Ls} \quad (5)$$

$$G_{ff}(s) = \frac{1}{1+\tau_{ff}s} \quad (6)$$

As it is represented by Norozpour Niazi et al. [8], the bode diagrams of the control scheme of Figure 1 should be assessed to regulate the parameters of the PR controllers. As it is presented there, to define these controller parameters which are located in internal and external loops and reach a reasonable proficiency, closed-loop stability of the system, and make small steady-state errors, the internal loop should be faster than the external one by defining smaller control system closed-loop bandwidth (near ten times smaller than switching frequency of the VSC). Similarly, the external loop should be slower than the internal loop, to prevent unstable control occur. In addition, the preferred tolerable range of phase margin is 30° to 60° [26].

2. 2. Primary Control Level The primary control level has two parts: the droop-based control and virtual impedance loop; to make accurate reference voltage by online creation of the frequency and amplitude of the voltage. Additionally, to achieve power with high reliability, the responses of primary control level to grid variation should be adequately fast.

2. 2. 1. Droop Control As is presented in Figure 2 and Equations (7)-(12) based on droop control, the sharing powers are attained by measuring the output current and voltage of the VSI [27]. Hence, according to instantaneous active and reactive powers, the amplitude and frequency of the reference signal are made. These calculated instantaneous powers have AC and DC components that necessitate the use of the LPF to make the fundamental-based powers extraction which is presented in Figure 2.

$$P = V_{c\alpha}I_{o\alpha} + V_{c\beta}I_{o\beta} \quad (7)$$

$$Q = V_{c\beta}I_{o\alpha} + V_{c\alpha}I_{o\beta} \quad (8)$$

$$f = f^* - G_p(s)(P - P^*) \quad (9)$$

$$E = E^* - G_q(s)(Q - Q^*) \quad (10)$$

$$G_p(s) = \frac{K_{Pp}s + K_{Ip}}{s} \quad (11)$$

$$G_q(s) = \frac{K_{Pq}s + K_{Iq}}{s} \quad (12)$$

where * represents the reference component

2. 2. 2. Selective Virtual Impedance Loop As is mentioned in the former research [8], to make better-decoupled power-sharing (P-Q), the SHVI loop has been proposed in this study, too. It should be mentioned that virtual resistance will cause an increase in the damping of a system without any losses on proficiency while virtual inductance will reduce power oscillation and current circulation among VSIs, increase the stability of the systems, and improve the act of decoupling in active and reactive power-sharing. Moreover, the harmonic voltage drops of lines and filter impedances can be compensated by virtual admittance [28]. Therefore, the primary current peak at the PCC is detracted and restricted by the output impedance which is increased through a virtual impedance. Hence, it makes simple DG plug-and-play capability and system stability [20]. Figures 3-4 and Equations (13)-(14) are represented that how the SHVI is added to the droop reference signal through MSOGI-FLL.

$$V_{v\alpha} = R^+_{v,f} I^+_{o\alpha,f} - \omega_f L^+_{v,f} I^+_{o\beta,f} + \sum_{h=1,-5,7,-11,13,\ldots} (R_{v,h} I_{o\alpha,h}) \quad (13)$$

$$V_{v\beta} = R^+_{v,f} I^+_{o\beta,f} + \omega_f L^+_{v,f} I^+_{o\alpha,f} + \sum_{h=1,-5,7,-11,13,\ldots} (R_{v,h} I_{o\beta,h}) \quad (14)$$

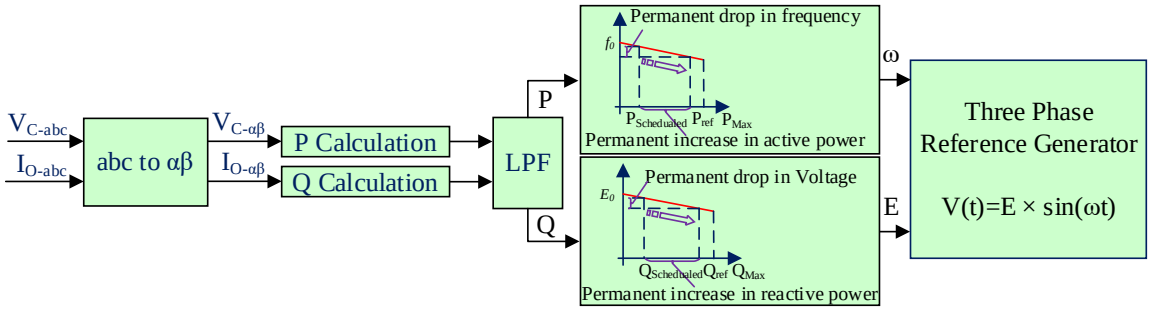


Figure 2. The block diagram of droop-based primary control scheme

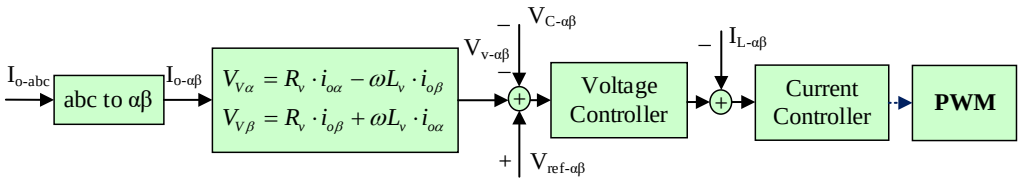


Figure 3. The block diagram of the virtual impedance loop (αβ frame).

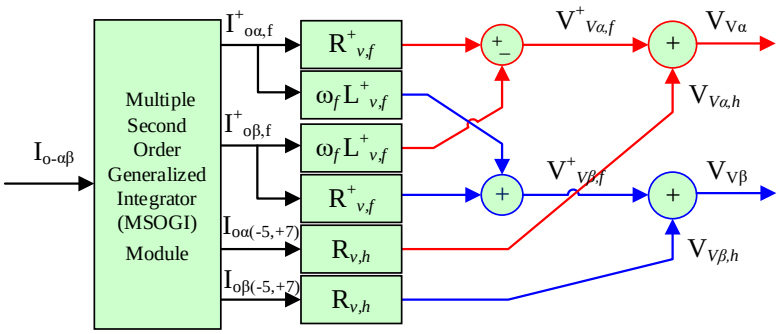


Figure 4. The block diagram of selective harmonic virtual impedance by MSOGI-FLL (αβ frame)

Moreover, it is denoted in [8, 29, 30] that to make extraction on the fundamental positive sequence and the harmonic components of the VSI output current, an

MSOGI-FLL is used. A typical block diagram MSOGI via a collection of parallel adaptive and selective filters which is tuned at the fundamental and different harmonic frequency in the frame of αβ is represented in Figure 5. Here, to find the fundamental frequency of the input current, a linked FLL to the SOGI-QSG is implemented, and to set the frequencies of them, the foreseeable frequency is applied with multiplying harmonic order to the frequency. Next, this structure is used to have different harmonic components detection of the input current [30].

2. 3. Secondary Control Level This level consists of two sections: synchronization and restoration; to attain zero steady-state error, frequency and voltage variation caused by the configuration of the grid, load changes, and droop controller.

Indeed, the restoration loop is required in the islanding mode operation of the MG to the frequency and voltage restoration. Moreover, to connect the DGs in MG and MG to the main grid, a synchronization loop is necessary [31].

2. 3. 1. Synchronization Control Loop in the Islanded MG To synchronize the DGs of MG and/or synchronize the microgrid with the upstream power system, a loop of synchronization is needed [32]. Here, to make the accurate frequency and voltage amplitude extraction of the energy resources and accordingly smooth transitions during the synchronization, the SRF-PLL with in-loop MAF is a good candidate for a distorted voltage which is discussed in our last study [8], too. As it is mentioned there, one of the abilities of the MAF is that the dc component is passed by them. The other capability is that it can block all harmonic components whose frequencies are integer multiples of $(1/T_{\omega})$. The SRF-PLL control loop with MAF has a high ability to make disturbance rejections. Figure 6 shows the block diagram of the synchronous reference frame phase locked-loop with in-loop MAF. In addition, the equations of frequency and voltage synchronization are as follows:

$$\omega_{syn} = \frac{K_{PF,syn}s + K_{IF,syn}}{s}(\omega^* - \omega_{SRF-PLL}) \quad (15)$$

$$E_{syn} = \frac{K_{PE,syn}s + K_{IE,syn}}{s}(E^* - E_{SRF-PLL}) \quad (16)$$

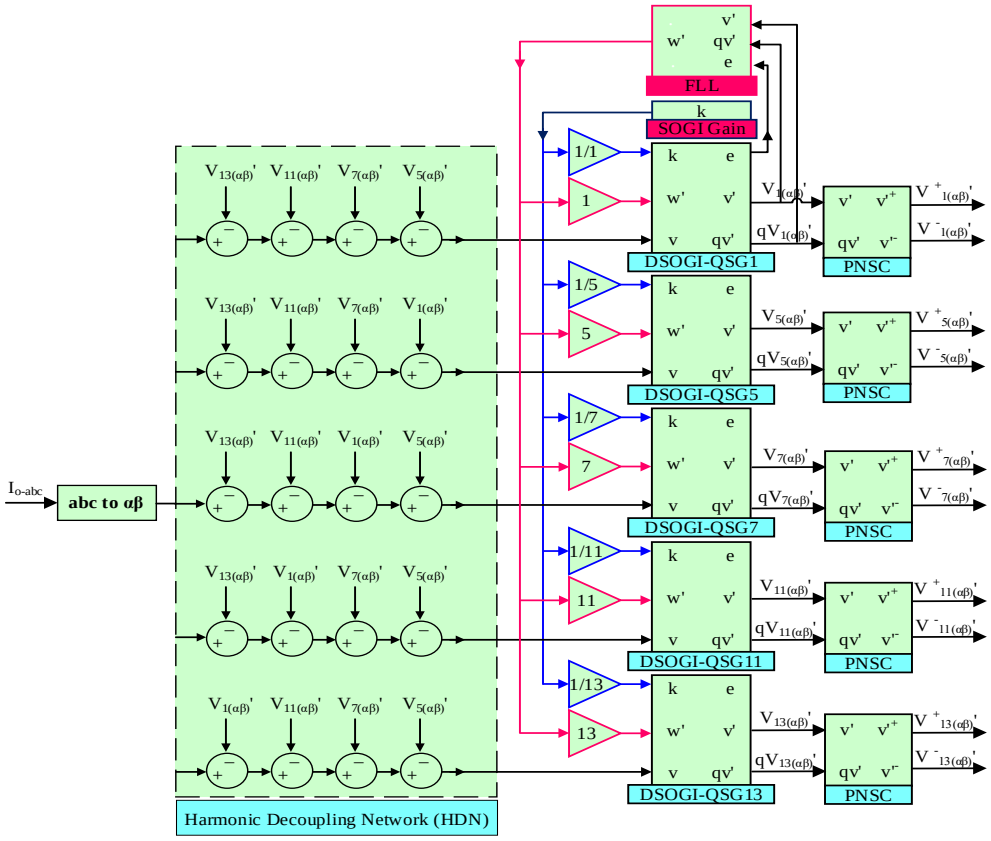


Figure 5. A typical block diagram of applied MSOGI (αβ frame)

2. 3. 2. Frequency and Voltage Restoration in the Islanded MG As is clear, one of the weaknesses of drop control is that it needs the frequency and voltage restoration to restore the voltage amplitude and frequency deviations to nominal values which are represented in Equations (17)-(18) and Figure 2.

$$\omega_{res} = \frac{K_{PF,res}s + K_{IF,res}}{s}(\omega^* - \omega_{SRF-PLL}) \quad (17)$$

$$E_{res} = \frac{K_{PE,res}s + K_{IE,res}}{s}(E^* - E_{SRF-PLL}) \quad (18)$$

It should be noted that the maximum frequency deviations based on the Nordel standard (North of Europe) and coordination of transmission of electricity union (Continental Europe) are 0.1 Hz and 0.3 Hz, respectively. Moreover, a maximum 10% deviation from the nominal value is defined for the voltage. Thus, the signal of restoration should be restricted to permissive frequency and amplitude of voltage deviations [20]. Figure 7 shows the implementation of restoration and synchronization signals in the secondary control level.

3. RESULTS AND DISCUSSION

In this study, a typical microgrid with three distributed resources, two nonlinear loads, and one linear load is simulated in the MATLAB/Simulink to assess the efficiency of the proposed hierarchical-based control structure and evaluate their transition behavior in grid-connected and islanding mode operations. The test system under analysis and simulations is presented in Figure 8. Here, as energy resources, three 2.5 kVA VSIs are employed which are powered by three 650-DC voltage sources. The parameters of MG DG units and loads are represented in Table 1.

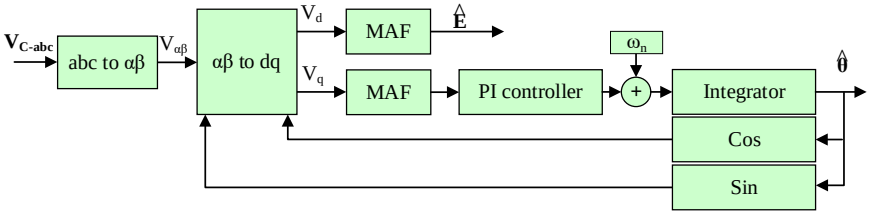


Figure 6. The SRF-PLL block diagram with in-loop MAF

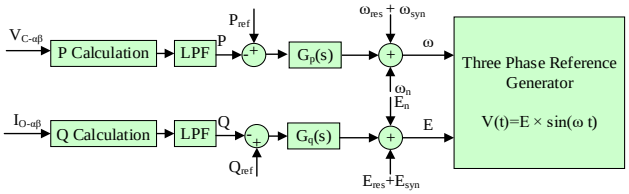


Figure 7. The restoration and synchronization block diagram in droop

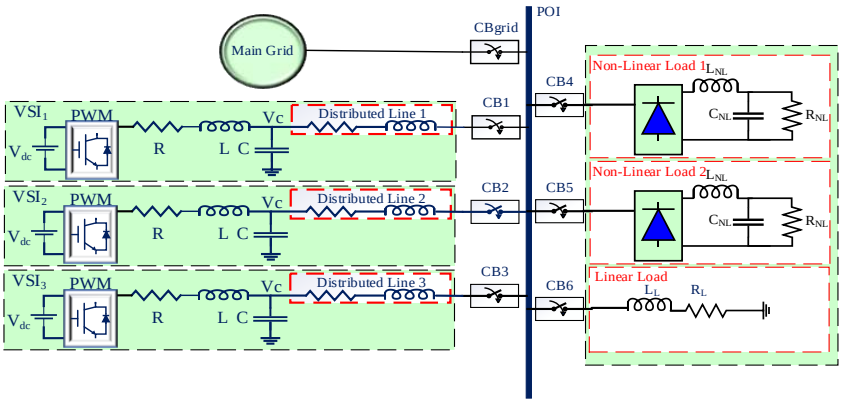


Figure 8. The test system under analysis and simulations

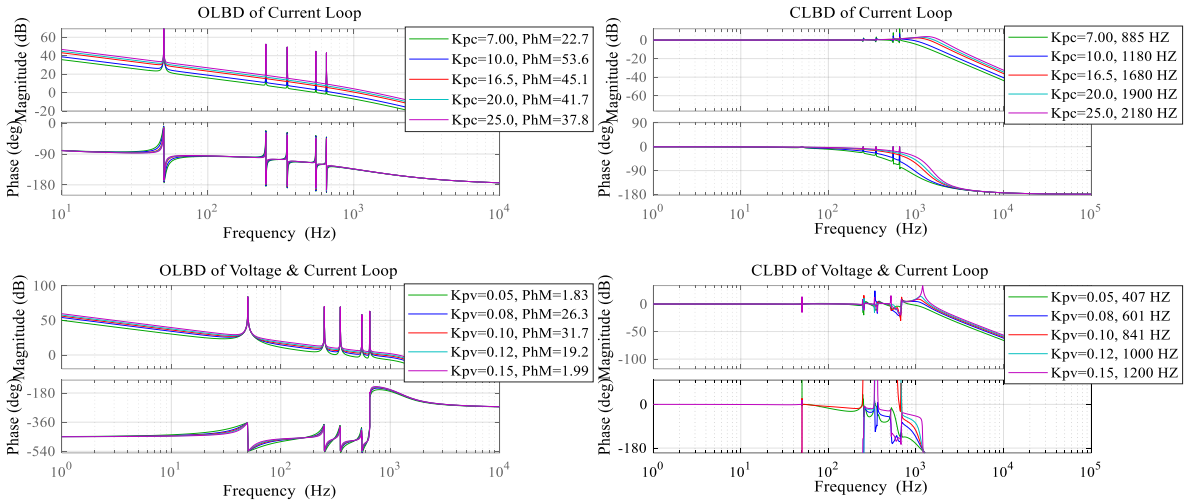


Figure 9. The bode plots of the internal and external loop of the inner level to design voltage and current controllers

TABLE 1. Load characteristics, power stage, and control parameters

LC Filters	R = 0.02 Ω, L = 1.8 mH, C = 25 µF
DC-Link Voltage	650V
Switching Frequency	10 kHz
DG Feeder	R_F = 0.2 Ω, L_F = 1.5 mH,
Main Grid (V/F)	400 V (line-line − RMS)/50 Hz
Linear Load	R_L= 200 Ω, L_L= 5 mH
Nonlinear Loads	R_{NL}= 100 Ω, L_{NL}= 0.84 mH, C_{NL}= 235 µF

Figure 9 is represented the bode plots of the internal loop and external loop to allocate voltage and current controller parameters. The behavior of variations of K_{pc} and K_{pv} on system phase stability margin in the internal and external loops are represented, there. By tunning the value of K_{pc}=16.5 and according to 10 kHz of the VSI switching frequency, the bandwidth and phase margin in the internal loop are achieved 1.68 kHz and 45.1°, respectively. These values can stable the current control system. As is mentioned before, to set the parameter of controllers and abstinence of interference, the internal loop should be faster than the external one. Then, by setting the value of K_{pv} in the external loop, 841 kHz for the bandwidth and 31.7° for the phase margin are achieved which can make a stable voltage control system. Besides, fundamental and harmonic orders (5^{th}, 7^{th}, 11^{th}, and 13^{th}) tracking will occur in both internal and external loops and approved that the controllers are well-analyzed to designed. Furthermore, Table 2 is presented the PR control parameters of internal and external loops. As the presence of nonlinear load leads to disturbed voltage, subsequently the input of standard SRF-PLL is a disordered signal.

TABLE 2. Proportional and multi-resonance current and voltage control parameters

$\mathbf{K_{pc}}$, $\mathbf{K_{rch1}}$, $\mathbf{K_{rch5}}$, $\mathbf{K_{rch7}}$, $\mathbf{K_{rch11}}$, $\mathbf{K_{rch13}}$	16.5, 200, 150, 150, 150, 150
$\mathbf{K_{pv}}$, $\mathbf{K_{rvh1}}$, $\mathbf{K_{rvh5}}$, $\mathbf{K_{rvh7}}$, $\mathbf{K_{rvh11}}$, $\mathbf{K_{rvh13}}$	0.10, 15.5, 15, 15, 15, 15

Therefore, on the base of Equations (19)-(21), the presence of $(h)^{th}$ order harmonic there causes $(h-1)^{th}$ order disturbance component in the control loop of PLL.

$$\begin{aligned} E_a(t) &= E_1\cos\varphi_f + E_h\cos\varphi_h \\ E_b(t) &= E_1\cos(\varphi_f - 2\pi/3) + E_h\cos(\varphi_h - 2\pi/3) \\ E_c(t) &= E_1\cos(\varphi_f + 2\pi/3) + E_h\cos(\varphi_h + 2\pi/3) \end{aligned} \quad (19)$$

$$E_d(t) = E_1\cos(\varphi_f - \varphi) + E_h\cos(\varphi_h - \varphi) \approx E_1 - E_h \quad (20)$$

$$\begin{aligned} E_q(t) &= E_1\sin(\varphi_f - \varphi) + E_h\sin(\varphi_h - \varphi) \\ &\approx E_1(\varphi_f - \varphi) - E_h(\varphi_h - \varphi) \end{aligned} \quad (21)$$

Therefore, harmonic orders of (−5, + 7, − 11, + 13, etc.), in the input of the SRF-PLL will appear as harmonic orders of (−6, + 6, − 12, + 12, etc.) in the PLL control loop. Likewise, the moving average filter which is a linear-phase low pass filter is applied in this study. Then, the DC component is passed and all harmonics whose frequencies are integer multiples of $1/T_w$ (T_w is the period of fundamental voltage) are entirely blocked. In the case of 50 Hz, harmonics up to the aliasing frequency are blocked (i.e. 50, 100, 150 Hz, etc). Furthermore, by considering T_w=T/2, all even-order harmonics up to the aliasing frequency (i.e. 100, 200, 300 Hz, etc) are blocked. Table 3 summarized the parameters of the SRF-PLL with in-loop MAF which is designed in our last study [8] and is modified and implemented, here.

TABLE 3. SRF-PLL within loop MAF parameters

T_w	0.02, S
$K_{p,\ PLL}$	82.8427
$K_{i,\ PLL}$	2.8427e+03, s^{-1}

TABLE 4. DGs Droop control parameters

K_{pp}, K_{ip}	1.142×10^{-4} (Ws/rd), 0.9×10^{-3}, (W/rd)
K_{pq}, K_{iq}	0.531×10^{-1} (VAR/V), 1.00 (VARs/V)

TABLE 5. Selective harmonic virtual impedance parameters

$R^{+}_{v,f}$, $L^{+}_{v,f}$, $R_{v,h}$ $(-5, 7, -11, 13)^{th}$	0.5 Ω, 0.01 H, 1.0 Ω

TABLE 6. Synchronization and restoration PI and adaptive PI controller parameters

$K_{PF,syn}$, $K_{IF,syn}$	0.0001, 10 s^{-1}
$K_{PE,syn}$, $K_{IE,syn}$	0.0015, 0.00001 s^{-1}
$K_{PF,sec}$, $K_{IF,sec}$	1.25, 0.83 s^{-1}
$K_{PE,sec}$, $K_{IE,sec}$	1.0, 0.075 s^{-1}
$K_{PF,Adap\text{-}sec}$, $K_{IF,Adap\text{-}sec}$	1.8, 4.83 s^{-1}
$K_{PE,Adap\text{-}sec}$, $K_{IE,Adap\text{-}sec}$	3.5, 10.075 s^{-1}

As shown in Figure 8, the capacitor inside the nonlinear loads can make an inrush current. Then, the transition in the time of system variations such as load connection/disconnection depends on its capacitor initial charge. Therefore, to limiting them which causes the inrush current of the capacitor, two-step series resistors are used for a short time (0.3 s). Here, two cases are studied to prove the proficiency of this research in MATLAB/Simulink.

Case study 1: three DGs are considered as MG and are responsible for injecting the active and reactive power to the power system and support local linear and nonlinear loads (as sensitive loads). In this case. it should be mentioned that the contribution of all DG units is considered equal. Consequently, the hierarchical-based control structure of three-phase VSIs in a type of MG is implemented to make a seamless transition in system variations and make harmonic compensation by proportional and multi-resonant voltage and current controllers. Also, SRF-PLL with in-loop MAF is used to prepare the V/F extraction of the VSIs and enhance the simulation results.

Case study 2: In this case, to consider the robustness of the hierarchical control more, it is assumed that the impedance line of DG_3 is twice the others. Also, the effect of implementing the adaptive secondary control in MG restoration is considered. Besides, to have better and decoupled harmonic power-sharing, selective harmonic virtual impedance is implemented.

As presented in Table 7, the simulation schedule is as follows: in T_1, the DGs synchronization of the VSIs are started. After DGs synchronization, the DGs are connected, and an MG is created. Then, MG synchronization starts in T_2 until T_3 to synchronize them to the main grid. In T_3, microgrid synchronization is finished and the microgrid is coupled to the upstream grid and is injected pre-planned power to them, simultaneously. Then, in T_4 and T_5, the local sensitive loads are linked, respectively. Soon after, unintended islanding of the MG has happened and the microgrid is responsible to support the local loads in T_6. Finally, in T_7, an unintentional outage occurs for DG_3 while other DGs are in control to support sensitive loads.

The simulation results of the frequency and maximum voltage amplitude waveforms of the VSIs in case 1 are represented in Figure 10. Likewise, Figure 11 is represented the controlled power-sharing of the DGs. As is shown there, all the DGs were injecting pre-planned power to the upstream grid (1350 W-450 VAR) by themselves at T_3 and after that, they support local sensitive loads which are linear and nonlinear at T_4 and T_5, respectively. Next, the MG is suddenly isolated at T_6, and accordingly, the exchanged power between the MG and the upstream power system is becoming zero. So, the MG should support local and sensitive loads, lonely. Also, an unintentional outage of DG_3 happens at T_7. In this case study, the parameters of all DGs and transmission lines are the same. In this paper, the advanced SRF-PLL with in-loop MAF which has higher disturbance rejection ability is employed, as well. It contains filters in the SRF-PLL to pass the DC component of the signals and filters the harmonic. Finally, as is showed in the achieved results of case 1, the seamless transitions from grid-connection mode to islanded mode, load variation, and distributed resource outage are attained. Plus, the harmonics compensation is implemented by the PR controller which is reported in Table 8. In case 2, to make more decoupled harmonic power-sharing in a system with nonlinear loads, a selective harmonic virtual impedance is implemented, too. Moreover, to cover the local grid variation in IM, a

TABLE 7. simulation schedule

Operation mode	Time (s)
time to start DGs synchronization	$T_1 = 0.0$
time to end of DGs synchronization and connection, time to start MG synchronization to the grid	$T_2 = 1.0$
time to end of MG synchronization, MG to grid connection, power injection to the main grid	$T_3 = 2.0$
time for connecting a linear local load	$T_4 = 3.5$
time for connecting nonlinear local loads 1, 2	$T_5 = 4.0$
time for unintentional islanding of MG	$T_6 = 5.0$
time for the unplanned outage of DG_3	$T_7 = 6.5$

simple adaptive secondary control is applied instead of general secondary control implemented in case 1 which is based on a two-step lookup table. Figure 12 is represented the measured voltage and frequency of the DGs by SRF-PLL with in-loop MAF. As is mentioned before, in this section, to consider the capability of the hierarchical control during system variation, it is assumed that the impedance line of DG_3 is twice the others. The effect of them is on DGs power-sharing. Figure 13 is shown the comparison of voltage and frequency in both cases 1 and 2. Although the behavior of control is acceptable in case 1, the control is more complicated in case 2. But the better seamless transition in voltage and frequency has happened which is shown in Figure 13.

Besides, Figure 14 is presented the decoupled active and reactive power shared by the DGs of the microgrid in case 2. As it is depicted there, in IM, although the droop coefficients of all DGs are the same, the reactive power is not shared equally between the DGs for the reason of the feeder line impedance mismatch and voltage drop in the feeder. Also, Figure 15 represents the injected current of the DGs to the PCC and main grid.

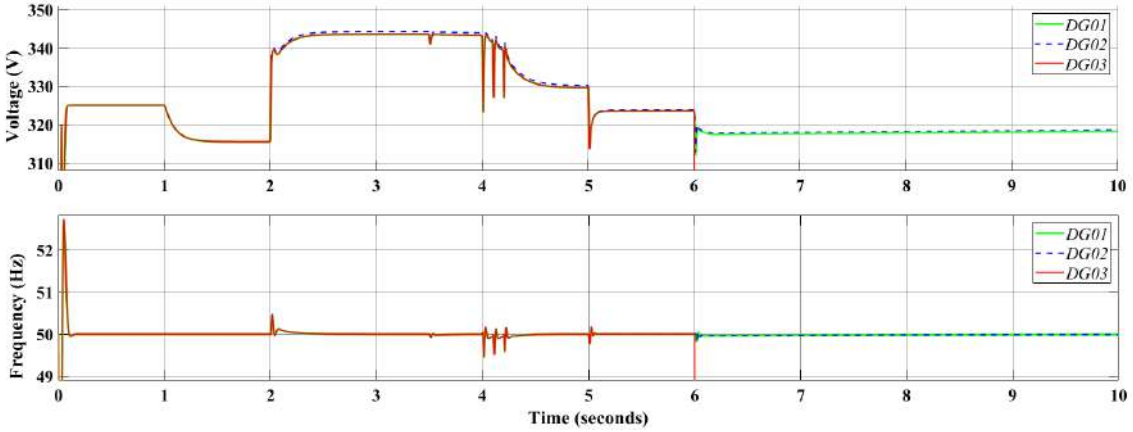


Figure 10. Simulation results of the frequency/voltage of the MG by SRF-PLL with in-loop MAF in Case 1

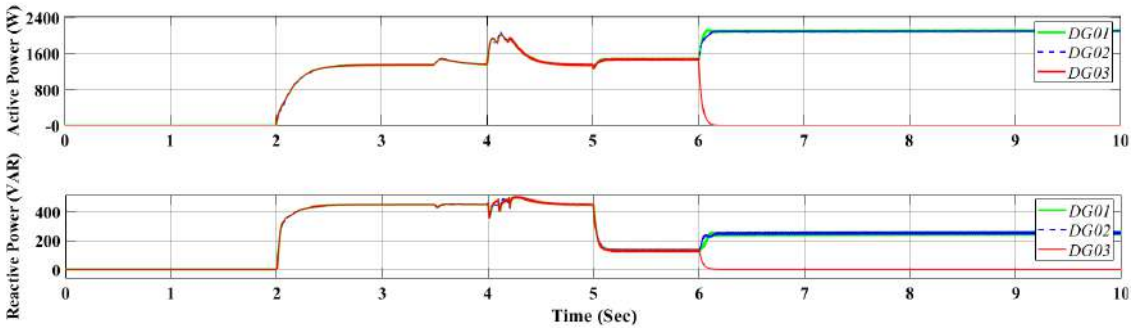


Figure 11. The controlled active and reactive power-sharing in the VSIs of the MG in case 1

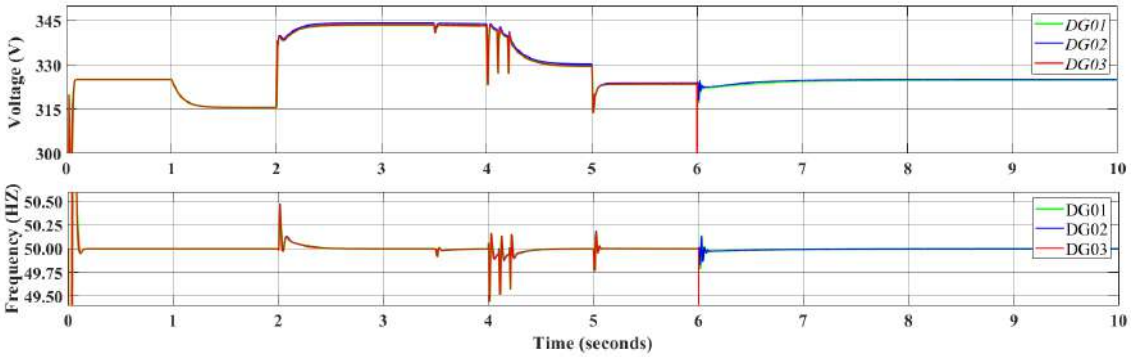


Figure 12. The simulation results of the frequency/voltage of the MG by SRF-PLL with in-loop MAF in Case 2

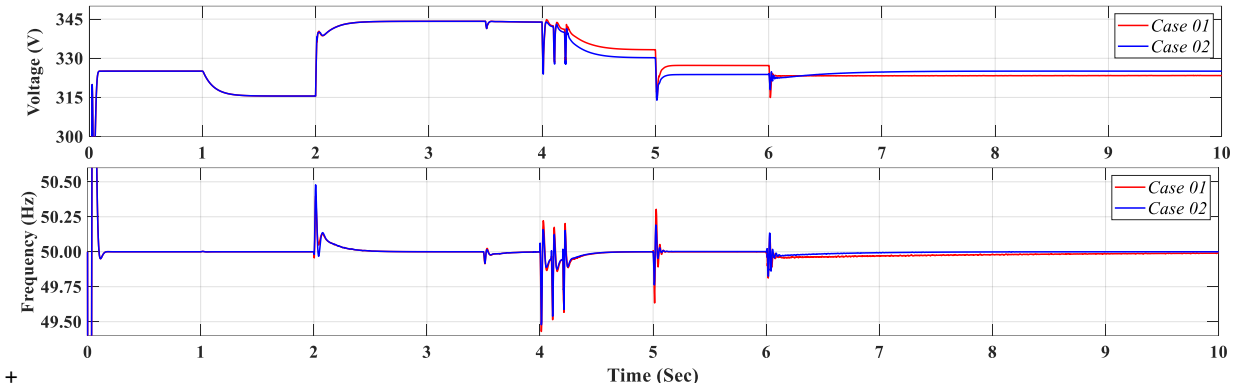


Figure 13. A comparison between different control approaches in voltage/frequency of the MG in Cases 1 and 2

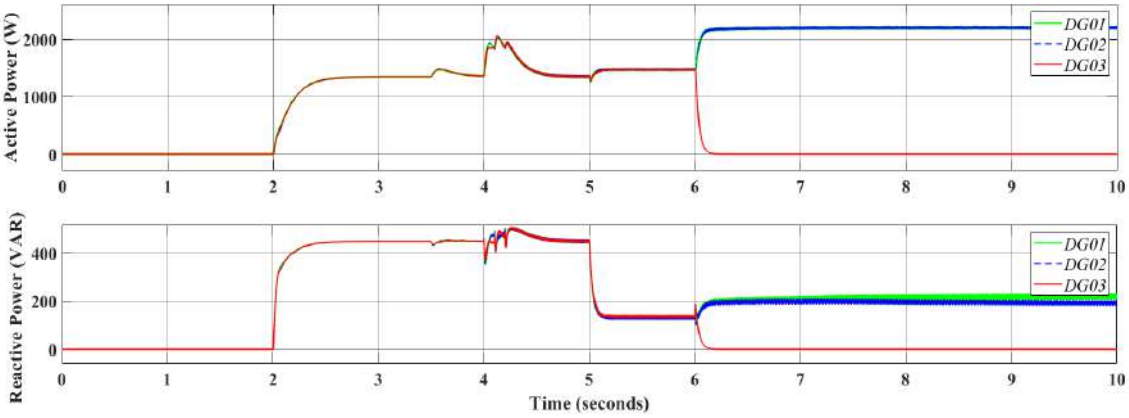


Figure 14. The controlled active and reactive power-sharing in the VSIs of the MG in case 2

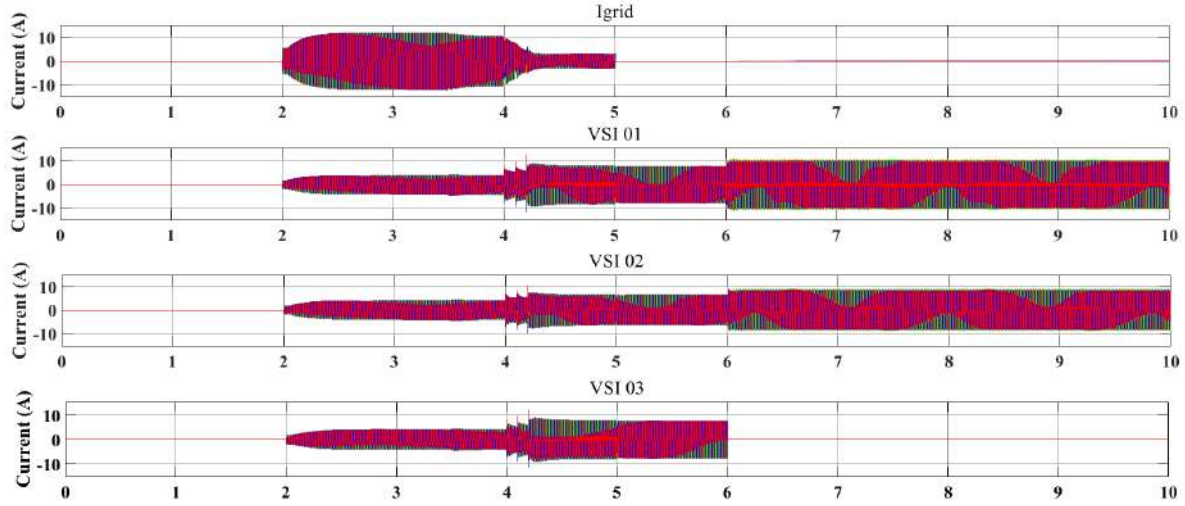
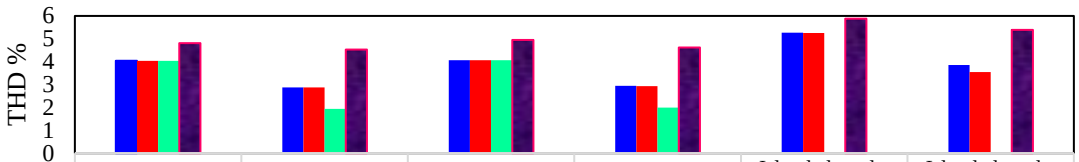


Figure 15. Three-phase injected current waveform to the PCC and main grid in case study 2

As the results are presented for both case studies in Figure 16, it is demonstrated that by applying SHVI in the primary control level of MG, hierarchical based control structure and adaptive secondary control, proper harmonic power-sharing and harmonic compensation along with seamless transition during grid variation are achieved.

4. CONCLUSION

This paper represented an enhanced control strategy based on hierarchy to enhance AC MG stability and make the seamless transition along with unintentional variation in the power systems. In the proposed structure, the PR controllers based on αβ frame are implemented to ensure



	Grid connected mode-Case 1	Grid connected mode-Case 2	Islanded mode-Case 1	Islanded mode-Case 2	Islanded mode, after DG3 outage-Case 1	Islanded mode, after DG3 outage-Case 2
VSI 1	4.04	2.87	4.07	2.94	5.27	3.85
VSI 2	4.03	2.87	4.06	2.93	5.25	3.54
VSI 3	4.03	1.94	4.06	1.99		
PCC	4.81	4.53	4.95	4.62	5.88	5.39

Figure 16. The THD (%) variation of the voltage in different operation modes

the stability and make compensation of harmonic while selective harmonic virtual impedance and MSOGI in the level of primary, simple adaptive lookup table restoration loop and advanced three-phase synchronous reference frame phase locked-loop with in-loop MAF in the level of secondary are corporating to damp variation and transition more. Here, the control performance is investigated by using grid dynamics such as unplanned islanding, loads dynamic, nonlinear loads existence, and DG's outages which may cause the insecure MG operation. Furthermore, an AC MG with interface three-phase PWM VSI is considered to confirm the efficacy of the planned control scheme and the simulation results denoted that the strategy has high performance in compensation of THD, robustness in controllability, seamless transition during variation of system configuration. The future work of this study will be extended to propose a novel and user-friendly controller based on delay control instead of the PR controller to control of microgrid.

5. REFERENCES

1. Palizban, O. and Kauhaniemi, K., "Hierarchical control structure in microgrids with distributed generation: Island and grid-connected mode," ***Renewable and Sustainable Energy Reviews,*** Vol. 44, 797-813, (2015), DOI: 10.1016/j.rser.2015.01.008.
2. Li, X., Zhang, H., Shadmand, M.B., and Balog, R. S., "Model predictive control of a voltage-source inverter with seamless transition between islanded and grid-connected operations," ***IEEE Transactions on Industrial Electronics,*** Vol. 64, No. 10, 7906-7918, (2017), DOI: 10.1109/TIE.2017.2696459.
3. Meng, L., Savaghebi, M., Andrade, F., Vasquez, J. C., Guerrero, J. M., and Graells, M., "Microgrid central controller development and hierarchical control implementation in the intelligent microgrid lab of Aalborg University," 2015 IEEE Applied Power Electronics Conference and Exposition (APEC), 2585-2592, (2015), DOI: 10.1109/APEC.2015.7104716.
4. Kaur, A., Kaushal, J., and Basak, P., "A review on microgrid central controller," ***Renewable and Sustainable Energy Reviews,*** Vol. 55, 338-345, (2016), DOI: 10.1016/j.rser.2015.10.141.
5. Guerrero, J. M., Chandorkar, M., Lee, T. and Loh, P. C., "Advanced Control Architectures for Intelligent Microgrids—Part I: Decentralized and Hierarchical Control," ***IEEE Transactions on Industrial Electronics,*** Vol. 60, No. 4, 1254-1262, (2013), DOI: 10.1109/TIE.2012.2194969.
6. Farokhian Firuzi, M., Roosta, A., and Gitizadeh, M., "Stability analysis and decentralized control of inverter-based ac microgrid," ***Protection and Control of Modern Power Systems,*** Vol. 4, (2019), 10.1186/s41601-019-0120-x.
7. Rocabert, J., Luna, A., Blaabjerg, F., and Rodriguez, P., "Control of power converters in AC microgrids," ***IEEE Transactions on Power Electronics,*** Vol. 27, No. 11, 4734-4749, (2012), DOI: 10.1109/TPEL.2012.2199334.
8. Norozpour Niazi, A., Vasegh, N., and Motie Birjandi, A. A., "To study unplanned islanding transient response of microgrid by implementing MSOGI and SRF-PLL based hierarchical control in the presence of nonlinear loads," ***IET Renewable Power Generation,*** Vol. 14, No. 5, 881-890, (2020), 10.1049/iet-rpg.2019.0506.
9. Guerrero, J. M., Vasquez, J. C., Matas, J., de Vicuna, L. G. and Castilla, M., "Hierarchical Control of Droop-Controlled AC and DC Microgrids: A General Approach Toward Standardization," ***IEEE Transactions on Industrial Electronics,*** Vol. 58, No. 1, 158-172, (2011), DOI: 10.1109/TIE.2010.2066534.
10. Sedhom, B. E., El-Saadawi, M. M., Hatata, A. Y., and Abd-Raboh, E. H. E., "H-infinity versus model predictive control methods for seamless transition between islanded- and grid-connected modes of microgrids," ***IET Renewable Power Generation***, Vol. 14, No. 5, 856-870, (2020), DOI: 10.1049/iet-rpg.2019.0018
11. Moradi,M. H., Eskandari, M., and Siano, P., "Safe transition from connection mode to islanding mode in microgrids," 24th Iranian Conf. on Electrical Engineering (ICEE), 1902-1907, (2016), DOI: 10.1109/IranianCEE.2016.7585832.
12. Li, Y., Yuan, L., Meng, K., and Dong, Z., "Smooth states transition control strategy for microgrid," IEEE Int. Conf. on Information and Automation (ICIA)**,** 86-91, (2017), DOI: 10.1109/ICInfA.2017.8078887.
13. Lavanya, V., and Senthil Kumar, N., "Seamless Transition in Grid-connected Microgrid System using Proportional Resonant Controller." ***International Journal of Engineering,*** Vol. 33, No. 10, 1951-1958, (2020), DOI: 10.5829/IJE.2020.33.10A.13.
14. Ahmadi, M., Sharafi, P., Mousavi, M. H., and Veysi, F., "Power Quality Improvement in Microgrids using STATCOM under Unbalanced Voltage Conditions." ***International Journal of Engineering, Transactions C: Aspects,*** Vol. 34, No. 6, 1455-1467, (2021), DOI: 10.5829/IJE.2021.34.06C.09.
15. Imran, R. M., Wang, S., Flaih, F. M. S, and Salih, H. W., "Smooth mode transition control for micro-grid with hybrid battery/diesel combination," 4th International Confernce on Information Science

and Control Engineering (ICISCE), 1237-1242, (2017), DOI: 10.1109/ICISCE.2017.257.

16. Golestan, S., Monfared, M., and Freijedo, F. D., "Design-oriented study of advanced synchronous reference frame phase-locked loops," ***IEEE Transactions on Power Electronics,*** Vol. 28, No. 2, 765-778, (2013), DOI: 10.1109/TPEL.2012.2204276.
17. Golestan, S., Guerrero, J. M., and Vasquez, J. C., "A PLL-based controller for three-phase grid-connected power converters," ***IEEE Transactions on Power Electronics,*** Vol. 33, No. 2, 911-916, (2018), DOI: 10.1109/TPEL.2017.2719285.
18. Golestan, S., Guerrero, J. M., Vidal, A., Yepes, A. G., and Doval-Gandoy, J., "PLL with MAF-based prefiltering stage: small-signal modeling and performance enhancement," ***IEEE Transactions on Power Electronics,*** Vol. 31, No. 6, 4013-4019, (2016), DOI: 10.1109/TPEL.2015.2508882.
19. Golestan, S., Guerrero, J. M., Vasquez, J. C., Abusorrah, A. M., and Al-Turki, Y., "A study on three-phase FLLs," ***IEEE Transactions on Power Electronics,*** Vol. 34, No.1, 213-224, (2019), DOI: 10.1109/TPEL.2018.2826068.
20. Vasquez, J. C., Guerrero, J. M., Savaghebi, M., Eloy-Garcia, J. and Teodorescu, R., "Modeling, Analysis, and Design of Stationary-Reference-Frame Droop-Controlled Parallel Three-Phase Voltage Source Inverters," ***IEEE Transactions on Industrial Electronics,*** Vol. 60, No. 4, 1271-1280, (2013), DOI: 10.1109/TIE.2012.2194951.
21. He, J., Li, Y. W., and Blaabjerg, F., "Flexible Microgrid Power Quality Enhancement Using Adaptive Hybrid Voltage and Current Controller," ***IEEE Transactions on Industrial Electronics,*** Vol. 61, No. 6, 2784-2794, (2014), DOI: 10.1109/TIE.2013.2276774.
22. Han, Y., Shen, P., Zhao, X., and Guerrero, J. M., "An enhanced power-sharing scheme for voltage unbalance and harmonics compensation in an islanded AC microgrid," ***IEEE Transactions on Energy Conversion,*** Vol. 31, No. 3, 1037–1050, (2016), DOI: 10.1109/TEC.2016.2552497.
23. Savaghebi, M., Jalilian, A., Vasquez, J. C., and Guerrero, J. M., "Secondary control for voltage quality enhancement in microgrids," ***IEEE Transactions on Smart Grid,*** Vol. 3, No.4, 1893-1902, (2012), DOI: 10.1109/TSG.2012.2205281.
24. Yazdani, A., Iravani, R., "Voltage-Sourced Converters in Power Systems: Modeling, Control, and Applications," ***Wiley-IEEE Press,*** 2010.
25. Rocabert, J., Luna, A., Blaabjerg, F., and Rodríguez, P., "Control of Power Converters in AC Microgrids," ***IEEE Transactions on Power Electronics,*** Vol. 27, No. 11, 4734–4749, (2012), DOI: 10.1109/TPEL.2012.2199334.
26. Zamora, R. and Srivastava, A. K., "Controls for microgrids with storage: review, challenges, and research needs," ***Renewable and Sustainable Energy Reviews.,*** Vol. 14, No. 7, 2009–2018, (2018), DOI: 10.1016/j.rser.2010.03.019.
27. Golsorkhi, M. S. and Savaghebi, M., "A Decentralized Control Strategy Based on V-I Droop for Enhancing Dynamics of Autonomous Hybrid AC/DC Microgrids," ***IEEE Transactions on Power Electronics,*** Vol. 36, No. 8, 9430-9440, (2021), DOI: 10.1109/TPEL.2021.3049813.
28. Mousazadeh Mousavi, S. Y., Jalilian, A., Savaghebi, M., and Guerrero, J. M., "Coordinated control of multifunctional inverters for voltage support and harmonic compensation in a grid-connected microgrid," ***Electric Power Systems Research,*** Vol. 155, 254-264, (2018), DOI: 10.1016/j.epsr.2017.10.016.
29. Golestan, S., Guerrero, J. M., Vasquez, J. C., Abusorrah, A. M., and Al-Turki, Y., "Modeling, Tuning, and Performance Comparison of Second-Order-Generalized-Integrator-Based FLLs," ***IEEE Transactions on Power Electronics,*** Vol. 33, No. 12, 10229-10239, (2018), DOI: 10.1109/TPEL.2018.2808246.
30. Rodriguez, P., Luna, A., Candela, I., Mujal, R., Teodorescu, R., and Blaabjerg, F., "Multiresonant Frequency-Locked Loop for Grid Synchronization of Power Converters Under Distorted Grid Conditions," ***IEEE Transactions on Industrial Electronics,*** Vol. 58, No. 1, 127-138, (2011), DOI: 10.1109/TIE.2010.2042420.
31. Heydari, R., Khayat, Y., Naderi, M., Anvari-Moghaddam, A., Dragicevic, T. and Blaabjerg, F., "A Decentralized Adaptive Control Method for Frequency Regulation and Power Sharing in Autonomous Microgrids," 2019 IEEE 28th International Symposium on Industrial Electronics (ISIE), 2427-2432, (2019), DOI: 10.1109/ISIE.2019.8781102.
32. Li, M., Gui, Y., Guan, Y., Matas, J., Guerrero, J. M. and Vasquez, J. C., "Inverter Parallelization for an Islanded Microgrid Using the Hopf Oscillator Controller Approach with Self-synchronization Capabilities," ***IEEE Transactions on Industrial Electronics,*** (2020), DOI: 10.1109/TIE.2020.3031520.

Persian Abstract

چکیده

در این پژوهش، کنترل سلسله مراتبی بهبود یافته ریزشبکه مبتنی بر کنترل‌کننده‌های تناسبی–چندگانه رزونانسی به منظور جبران‌سازی اغتشاشات هارمونیکی ناشی از بارهای غیرخطی پیشنهاد شده است. همچنین گذارهای محتمل ریزشبکه، بویژه از حالات متصل به شبکه بالادست به حالت جزیره‌ای همراه با گذارهای برخواسته از خروج‌های ناخواسته ریزمنابع موجود در یک ریزشبکه مورد مطالعه قرار گرفته است. در اینجا، ساختار کنترل سلسله مراتبی پیشنهاد شده از سه سطح کنترلی تشکیل شده که شامل سطوح کنترلی داخلی، اولیه و ثانویه بر مبنای قاب مرجع ساکن می‌باشند. کنترل‌کننده‌های ولتاژ و جریان اینورترهای مبدل ولتاژی سه فاز که در سطح داخلی قرار گرفته‌اند، از نوع کنترل‌کننده های تناسبی– چندگانه رزونانسی می‌باشند. بعلاوه در سطح کنترل اولیه، از یک کنترل‌کننده اُفتی همراه با امپدانس مجازی هارمونیکی گزینشی برای رسیدن به تسهیم مناسب توان بین ریزمنابع استفاده شده است. به منظور رسیدن به بازیابی بهتر و گذار هموارتر در جزیره‌ای شدن‌های برنامه‌ریزی نشده، خروج‌های ناخواسته ریزمنابع موجود در یک ریزشبکه و فرایند همگام‌سازی، از یک فاز حلقه قفل شده توسعه یافته در قاب مرجع ساکن همراه با حلقه فیلتر میانگذر متحرک و یک کنترل کننده تطبیقی بر مبنای جدول مرجع ساده در سطح کنترلی ثانویه بهره گرفته شده است. نتایج شبیه‌سازی در محیط متلب/سیمولینک نشان داده که روش پیشنهاد داده شده بازدهی سیستم کنترل، موثر بودن آن و مقاوم بودن آن را در حین تغییرات در ساختار شبکه بهبود داده است.

International Journal of Engineering

Journal Homepage: www.ije.ir

Memristor Crossbar-based Hardware Implementation of Type-2 Fuzzy Membership Function and On-chip Learning

S. Haghzad Klidbary*[a], M. Javadian[b], R. Omidi[c], R. P. R. Hasanzadeh[d]

[a] *Department of Electrical and Computer Engineering, University of Zanjan, Zanjan, Iran*
[b] *Computer Department, Kermanshah University of Technology, Kermanshah, Iran*
[c] *Department of Electrical Engineering, University of Zanjan, Zanjan, Iran*
[d] *Department of Electrical Engineering, University of Guilan, Rasht, Iran*

PAPER INFO

Paper history:
Received 7 March 2021
Received in revised from 7 August 2021
Accepted 9 August 2021

Keywords:
Fuzzy Inference System
Fuzzy Membership Function
Type-1 Fuzzy
Type-2 Fuzzy
Hardware Implementation
Memristor-crossbar Structures

ABSTRACT

Utilizing fuzzy techniques, especially fuzzy type-2, is one of the most widely used methods in machine learning to model uncertainty. In addition to algorithm provision, the hardware implementation capability, and proper performance in real-time applications are other challenges. The use of hardware platforms that have biological similarities and are comparable to human neural systems in terms of implementation volume has always been considered. Memristor is one of the emerging elements for the implementation of fuzzy logic based algorithms. In this element, by providing current and selecting the appropriate direction for the applied current, the resistance of the memristor (memristance) will increase or decrease. Various implementations of type-1 fuzzy (T1F) systems exist, but no implementation of type-2 fuzzy (T2F) systems has been done based on memristors. In this paper, memristor-crossbar structures are used to implement type-2 fuzzy membership functions. In the proposed hardware, the membership functions can have any shape and resolution. Our proposed implementation of type-2 fuzzy membership function has the potential to learn (On-Chip learning capability regardless of the host system). Besides, the proposed hardware is analog and can be used as a basis in the construction of evolutionary systems. Furthermore, the proposed approach is applied to memristor emulator to demonstrate its correct operation.

***doi**: 10.5829/ije.2021.34.09C.15*

1. INTRODUCTION

Nowadays, the design of artificial intelligence-based computing machines that can intelligently perform human-like tasks is one of the main goals of researchers. Therefore, various algorithms have been proposed. The proposed algorithms can be categorized into the artificial neural networks, evolutionary algorithms and fuzzy systems. Neural Networks seek to find structures similar to the neural structures of living organisms, while the evolutionary algorithms try to seek swarm intelligence in nature. On the other hand, the fuzzy systems tries to find functional concepts identical to the human brain.

Fuzzy logic theory can be considered as a human effort to create a tool to deal with the inherent inaccuracy of systems. "Fuzzy logic" was first introduced in 1965. The word fuzzy means inaccurate, vague, and ambiguous. The idea of fuzzy sets was based on linguistic variables. In fuzzy logic, the problem is modeled as a group of "IF-THEN" rules, unlike the classical methods which model the problem by expressing complex mathematical relations. Fuzzy logic was utilized in various applications [1-5]]. When the efficiency of type-1 fuzzy (T1F) sets are approved in many research papers, Zadeh et al. [3] presented the second series of fuzzy sets. In this set, each degree of membership was also a member of a fuzzy set and named it the type-2 fuzzy (T2F) set. The T2F set was introduced in 1975 and is an extension of T1F sets. Some of the reasons for creating the T2F set are stated below [6, 7]: (1) The inability to

*Corresponding Author Institutional Email: s.haghzad@znu.ac.ir (S. H. Klidbary)

Please cite this article as: S. Haghzad Klidbary, M. Javadian, R. Omidi, R. P. R. Hasanzadeh, Memristor Crossbar-based Hardware Implementation of Type-2 Fuzzy Membership Function and On-chip Learning, International Journal of Engineering, Transactions C: Aspects, Vol. 34, No. 09, (2021) 2180-2188

mathematically describe systems in which the data-generating is known to be time-varying (mobile telecommunications); (2) The inability in the mathematical description of non-stationary noise; (3) Inability in the mathematical description of features in pattern recognition applications that have non-stationary statistical properties; (4) Knowledge obtained from a group of experts (expertise extraction) those using questionnaires that involve uncertain words; (5) Use linguistic expressions that have a non-measurable domain.

Currently, the T2F set has found many applications in the fields with high uncertainty factor. These applications include pre-processing of radiographic images [8], estimating the health of the newborn [9], controlling mobile robots [10], edge detection in digital photographs [11]. According to the increasing application of T2F sets in real-world engineering problems, there is a need for a fast hardware for implementation. Therefore, providing an efficient hardware for implementation of these algorithms would be beneficial for engineering problems. Because, the advantages of the soft computing algorithms will be observable when they can be used as hardware in the form of an intelligent system with high computational power and efficiency.

In hardware implementation, inspiration from nature and the implementation of brain-like structures have received much attention. The brain is structurally composed of a large number of neurons and synapses. It can be concluded that the density of neurons in the brain is very high by considering the dimensions of the brain. So far, many computational structures were developed based on CMOS technology [12-14]. Since this technology faces challenges at the nanoscale, it is not a proper option for implementing Bio-Inspired systems and achieving high volumes of processing elements. Therefore, it can be said that the manufacturing of more advanced technologies, more processing power, more compact circuit, and less power consumption are the most critical issues in the construction of an artificial brain. After the implementation of the memristor in HP (Hewlett Packard) Lab., the hope for achieving a computing platform comparable to the neural system of living organisms has increased.

Memristor is a nonlinear passive electrical element whose resistance is controlled by voltage. This element, by having a nanoscale size, small volume, low power consumption, and also identical behavior similar to the biological synapse, has been able to find an appropriate position in the hardware implementation of computational algorithms. This element attracted the attention of many researchers in various fields such as artificial neural networks and fuzzy systems. In this paper, the membership functions of T2F sets are implemented on the memristor-crossbar structure. Also, a circuit to type-reduction in this T2F set is given.

The rest of this paper is organized as follows. The main concepts of memristor element are reviewed in section 2. In section 3, the T2F set is discussed. The memristor-crossbar hardware implementation is presented in section 4. In section 5, the simulation results are reported. Finally, section 6 includes summaries, conclusions, and suggestions for future work.

2. MEMRISTOR BASICS

In addition to the three previously known fundamental electrical elements: capacitors, inductors, and resistance, in 1976, Chua and kang [22] introduced the fourth element by mathematical proof. This element, as a passive element, is the relationship between magnetic flux and electric charge [15, 16]. Chua [23] called this element as memristor, which stands for memory and resistor. Before 2008, no successful implementation of this element was reported, where the main reason was that the memristor property is visible in nanoscale. This element was successfully realized in mid-2008 [17]. After that, many researches have been done about memristor and its applications [18-21]. Memristor acts like a linear resistor, except that it has memory. This means that the memristor's properties at any given time depend on how the current or the voltage has been applied to it in the past. Figure 1 shows the relationships between different electrical parameters.

The feature of the memristor is that if the current passes through the memristor in one direction, its resistance will increase and if the direction of the current changes, the resistance value will decrease. So the memristor is called a memory resistor. In the following, the description of the memristor, which is provided by HP, is given by Chua [23].

2. 1. The Memristor Modeling This memristor is comprised of a very thin TiO_2 with the width, D, sandwiched between two platinum electrodes. The

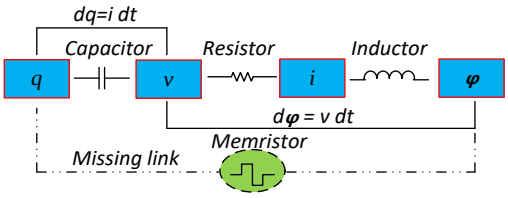


Figure 1. A representation of the four basic electrical elements, which are the resistor, the capacitor, the inductor, and the memristor. Memristor is the fourth fundamental element; the missing link represents the relationship between magnetic flux and electric charge. Theoretically, the information stored in memristors will be stored indefinitely if no current passes through it. Unlike a capacitor, it will not need to be refreshed

semiconductor layer has two regions, one with oxygen impurities and R_{on} resistance and the second without impurities with R_{off} resistance. By applying an external voltage, the border between the doped and undoped regions can be displaced (Figure 2(a)). This will change the memristance value of the memristor.

Memristance value varies between R_{on} and R_{off}. The memristance value reaches its minimum (equal to R_{on}) when the doped region extends to D ($w \to D$). The memristance value reaches its maximum (equal to R_{off}) when the undoped region extends to D ($w \to 0$). The mathematical model of the memristor, assuming the uniformity of the electric field inside the memristor, is as follows:

$$R_M(w) = R_{on} \times \frac{w(t)}{D} + R_{off} \times \left(1 - \frac{w(t)}{D}\right) \quad (1)$$

$$w(t) = w_0 + \frac{\mu_v R_{on}}{D} q(t) \quad (2)$$

where W_0 is the initial width of the doped region w, μ_v is ion mobility and $q(t)$ is the electrical charge that passes through the element.

A crucial point in a memristor is that memristor properties are visible in nanoscale. This characteristic of memristor makes it possible to implement hardware similar to the brain, which has a high density of neurons and synapses. The ability to store data in analog form, non-volatile memory, nanoscale fabrication and higher speed due to the nature of analog are some of the advantages of this element. Using the memristor alone is not optimal, and such as mentioned in Figure 3, this element is used as a crossbar structure. A crossbar consists of a series of horizontal wires passing over vertical ones. A memristor is placed at the intersection of two wires that cross each other. By applying proper voltage over any pair of vertical and horizontal wires, the memristor at that intersection can be accessed. Some implementations of fuzzy systems on memristor structures are given below. The memory property of memristor has been used, and by utilizing this feature, the four primary operations (addition, subtraction, multiplication, and division) have been implemented by Merrikh-Bayat and Shouraki [22]. Amer et al. [23] discussed on the implementation of Center Of Gravity (COG). The combination of memristor and resistor is used to make two operators, minimum and maximum [24]. The implementation of T1F membership functions on memristor-crossbar structures is given by Merrikh-Bayat et al. [25]. Merrikh-Bayat and Shouraki [26] presented, a new neuro-fuzzy computational structure. The implementation of ink drop spread (IDS) in 2D plans on the memristor-crossbar structure is presented by Merrikh-Bayat et al. [27]. There are also other references related to the implementation of the ALM (Active Learning Method) algorithm on memristor-crossbar [28-30]. Memristor has also been used as a synapse in neuromorphic computing [31-34]. Figure 3 shows the memristor-crossbar structure and its applications as a synapse is given.

As we know, fuzzy systems can be considered as human efforts to create a tool to deal with the inherent inaccuracies of systems. T2F set is an extension of the T1F set which has been able to provide proper results due to its ability to deal with the uncertainty of system rules and parameters. According to the researches, it can be stated that so far, in addition to all the implemented cases expressed on the memristor-crossbar structures, no implementations of T2F membership functions have been observed on the memristor structure. In the following, we will first give an overview of the T2F set.

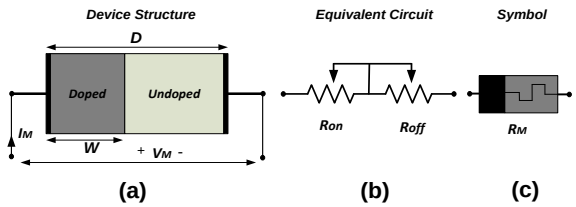


Figure 2. (a) the structure of the memristor, which consists of two regions respectively doped (low resistanceR_{on}) and undoped (high resistanceR_{off}). By applying a voltage (V_m) with appropriate polarity, this border can be changed, (b) circuit equivalent of memristor, (c) The symbol used for the memristor with R_M memristance

3. TYPE-2 FUZZY(T2F) SET

In the T1F set, a certain degree of membership between zero and one is considered. This specificity of the degree of membership challenges the uncertainty. After the aforementioned problems, the T2F set was proposed. The T2F set acts as a probability density function and provides uncertainty about the degree of membership using the variance around that degree of membership. Therefore, it can be mentioned that the T1F set models the first-degree of uncertainty, and the T2F set models the second-degree of uncertainty. Therefore, it can be

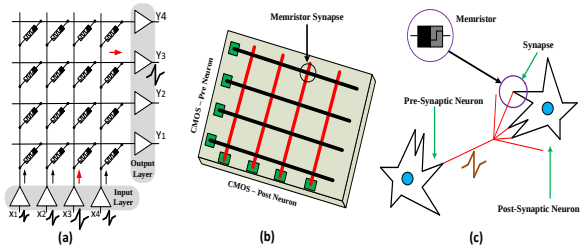


Figure 3. (a) Memristor-crossbar based circuit as a hardware implementation of spiky neural network (SNN), (b) the memristor-crossbar structure, (c) biological synapses

mentioned, the T2F system is also resistant to uncertainties that arise in fuzzy rules or system parameters. Equation (3) is the T1F set and Equations (4) and (5) present T2F set.

$$A = \sum_i \mu_A(x_i)/_{x_i} \ , x_i \in X \tag{3}$$

$$\tilde{A} = \{(x, \mu_{\tilde{A}}(x)) | \forall\, x \in X\} \tag{4}$$

$$\mu_{\tilde{A}}(x) = \sum_i f_x(u_i)/_{u_i} \ , \ u_i \in J_{x_i} \ , J_{x_i} \subseteq [0\,,1] \tag{5}$$

where x is primary variable, u is secondary variable, $\mu_{\tilde{A}}(x)$ presents the secondary membership function and $J_{x_1} \subseteq [0,1]$ is a set of primary membership degrees. T2F membership functions are described by a characteristic called FOU (footprints of uncertainty), which is a finite range of uncertainties and an expression of the union of primary membership functions. The FOU is enclosed between two T1F membership functions, which in Figure 4, the uncertainty effect by upper membership function (UMF) and lower membership function (LMF) is shown. It is worth mentioning that the values that secondary membership functions take are called the degree of secondary membership. FOU is utilized to eliminate $3D$ display problem of a T2F set that displays borders in a two-dimensional form. Figure 4(a) shows the T1F set, Figure 4(b) shows the T1F set without uncertainty, Figure 4(c) shows the interval T2F membership function and Figure 4(d) indicates the general T2F membership function. Commonly used methods such as Point-Valued Representation (PVR), Vertical-Slice Representation (VSR), Wavy Slice Representation (WSR) or Embedded T2FS Representation, and Horizontal Slice Representation (HSR) or α-plane Representation can be mentioned to express T2F sets. Computations in the T2F set require an algebraic operation and its dedicated fuzzy inference system, just like the T1F set. However, there are extensive researches in this field [35, 36].

After the inference operation, the output is a T2F set, and finally, the output must be brought to the crisp space. In the T2F set, like the T1F set, we need to use the defuzzification procedure. The mapping of a T2F set to a T1F set is called type reduction. The type reduction methods are the same as the methods developed in type-1 systems.

Among the most important methods of type reduction, Generalized Centroid and Karnik-Mendel (KM) can be mentioned. The Generalized Centroid method is the first method where in this method each embedded set has a center of gravity. In the following Equation, the calculation of this method is given and this Equation is calculated for each Embedded Set.

$$C_{\tilde{A}} = \left\{ \left(\frac{\sum_{i=1}^{N} x_i.u_i}{\sum_{i=1}^{N} u_i}, \mu_{\tilde{A}}(x_1)(u_1) * \ldots * \mu_{\tilde{A}}(x_N)(u_N) \right) | \forall i \in \{1, \ldots, N\} : x_i X\, u_i\, J_{x_i} \subseteq U \right\}, \quad t-norm(*) \tag{6}$$

The T-Norm operator is usually chosen to be minimal so that the membership degree obtained is not too small. Of course, by increasing the number of embedded sets, the calculations increase significantly. Another method is Vertical Slice Representation (Vertical Slice Centroid) and the focus of this paper is on this method. In this method, de-fuzzyization takes place in each vertical section, and finally a T1F set is obtained, which can be easily converted to a crisp number. The relation of this method are simple and have been avoided due to brevity.

In the following, the proposed memristor-crossbar hardware for implementation of the T2F membership function set and the type reduction section will be presented.

4. PROPOSED HARDWARE

As mentioned in the section 3, the memristor is a flexible circuit element whose resistance can be controlled by the current flow. In this paper, memristor-crossbar is used as a hardware platform for implementation due to their small size and high density in small volume compared to other elements and also the low energy required for changing the memristance value.

Utilizing memristor-crossbar structures has two phases (or states) of writing and reading. In the writing phase, information is stored on the memristor, which is called the learning phase. In the reading phase, which is called the usage or test phase, the stored information will be read. Intermediate circuits perform these phases on memristor-crossbar. The proposed hardware consists of four sections: control, memristor-crossbar, type-reduction (from T2F set to T1F set), and defuzzification.

The control section can be a microprocessor or FPGA or DSP, which is responsible for controlling the timing

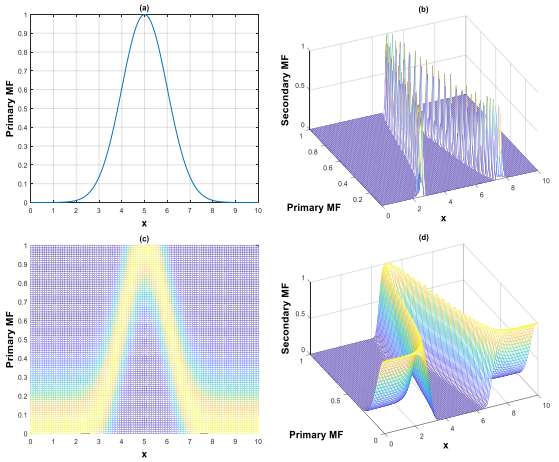


Figure 4. (a) type-1 membership function, (b) type-2 membership function without uncertainty, (c) interval T2F membership function and (d) general T2F membership function

and order of the applied signals. The two sections of type-reduction and defuzzification circuits are composed of resistors, Op-Amps, diodes, and transistors. In this section, we focus more on the memristor-based structure (for the simplicity, a 4x4 memristor-crossbar structure, as illustrated in Figure 5 is used).

For each type-2 function, there is a memristor-crossbar with R_x rows and R_u columns (R_x and R_u are the quantization levels). Each memristor is equivalent to one memory (pixel). These memristors are selected using de-multiplexers. In the training phase, a pulse is applied to the desired column, and the row corresponding to the value u is grounded as shown in Figure 6. At the training phase, the signal p is zero, and all the transistors connecting the rows of crossbar to the Op-Amps are disconnected, which causes the output to be zero volts at the beginning of the training (so these transistors act as isolators).

It can be mentioned that by using the ladder structure [28, 29], changes can be made as a Gaussian function in the memristance of memristors. As mentioned in section 3, Vertical Slice Representation (Vertical Slice Centroid) for T2F set is considered here. In this method, de-fuzzyization takes place in each vertical section, and finally a T1F set is obtained. In the type-reduction section, the maximum function is used. The circuit of this function is shown in Figure 7.

According to Figure 7, there are two series Op-Amps per input line, which is the same number for a system with R_u quantization level. In the test phase, the signal p is one then by applying V_{Read} to the desired rows and columns of the memristor-crossbar, the output voltage of each line is calculated as follows:

$$v_{OP1} = -V_{Read} \times \frac{R_f}{R_{m_{ij}}} \quad (7)$$

$$v_{o1} = V_{Bias} + v_{OP1} \quad (8)$$

where $R_{m_{ij}}$ indicates the memristance value of the memristor in the i_{th} row and the j_{th} column. Assuming the initial memristance values of all memristors are R_{on} and the values $R_f = R_{on}$, and also assuming $V_{Read} = V_{Bias}$, the output values are zero when no training has been done, which specifies the initial condition. In the learning process, memristance of memristors may increase, after which the absolute value of $-V_{Read}(\frac{R_f}{R_{m_{ij}}})$ begins to decrease, resulting in larger v_{o1} values. Therefore, for the first column ($x = 1$), the output voltage will be obtained as follows:

$$v_1 = max(V_{o1}, \dots, V_{o4}) - V_{Don} \quad (9)$$

Using the same hardware as Figure 7 for all columns, voltages will be obtained for each input. We can also use hardware such as weighted average (WA) instead of the connector block as shown in Figure 7.

4. 1. The Defuzzification Stage After type-reduction, the resulting outputs (voltage) indicate the membership degree in the corresponding inputs. The output is a vector, which is considered as a fuzzy number, so the output should be transformed into a crisp number

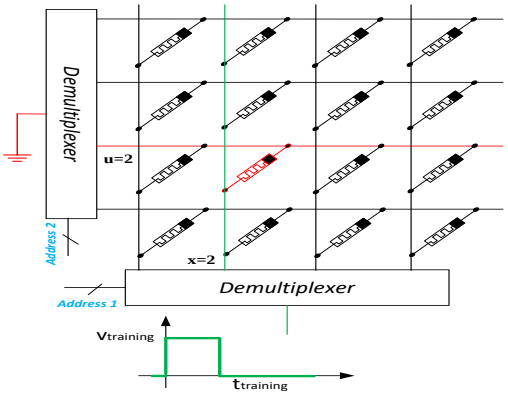


Figure 6. Selecting the desired memristor using de-multiplexers while applying a learning pulse (assumption of training for the second row ($u = 2$) and second column ($x = 2$))

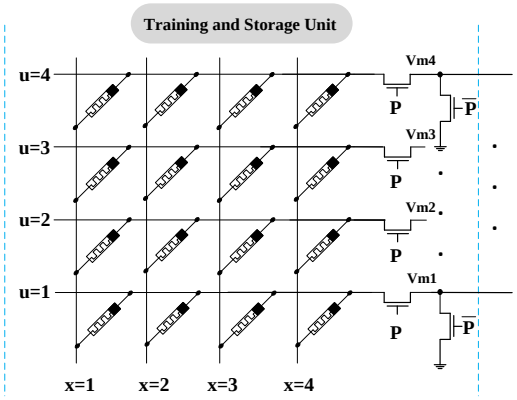


Figure 5. Proposed memristor-crossbar hardware for storing, learning, and testing operations

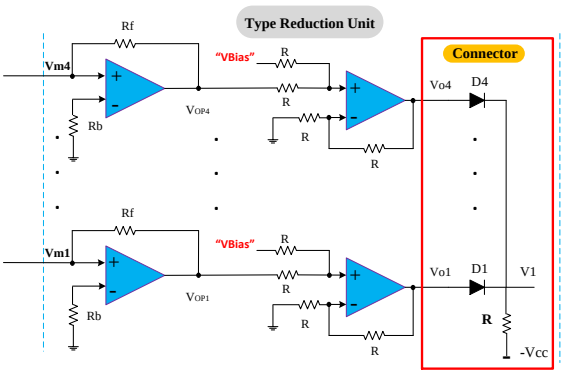


Figure 7. The type-reduction circuit, assuming the output v_1 for this circuit is $V_1 = max(V_{o1}, \dots, V_{o4}) - V_{Don}$

by an algorithm such as a WA. To implement the WA section as the defuzzification section, the hardware circuit that is shown in Figure 8 is used. The outputs v_{w1} and v_{w2} are calculated for n pairs $(x_i, \mu(x_i))$ according to the following equations.

$$v_{w1} = -(\mu_1 \times \frac{n\times R}{\frac{n}{1}\times R} + \mu_2 \times \frac{n\times R}{\frac{n}{2}\times R} + \cdots + \mu_n \times \frac{n\times R}{\frac{n}{n}\times R}) = -(\mu_1 + \mu_2 \times 2 + \cdots + \mu_n \times n) = -\sum_{i=1}^{n} \mu_i \times i = -\sum_{i=1}^{n} \mu_i \times x_i \quad (10)$$

$$v_{w2} = -(\mu_1 + \mu_2 + \cdots + \mu_n) = -\sum_{i=1}^{n} \mu_i \quad (11)$$

The obtained outputs are applied into an analog voltage divider, and the output is obtained according to the following equation.

$$v_{wA} = -\sum_{i=1}^{n} \mu_i \times x_i \,/ -\sum_{i=1}^{n} \mu_i \quad (12)$$

Voltage v_{wA} will be a number between 1 and the quantization level (n) of the input. This voltage can be converted to an analog voltage using a digital to analog converter (DAC).

It should be noted, the number of memristors can be increased in order to improve the accuracy, i.e., by increasing the number of quantization levels, the accuracy can be improved. In this implementation, the values are coded to memristance of memristors, so any desired value can be programmed by applying the appropriate voltage on memristors.

5. SIMULATION RESULTS

For investigating the accuracy of the proposed hardware, we conducted a T2F as shown in Figure 9. We have simulated the proposed circuit in HSPICE 2009 with TSMC 350nm technology. In this simulation, the memristor model proposed by Biolek et al. [37] has been used. Before simulating this hypothetical function, the first step is to examine the performance of each part of the circuit. One of the most important elements in the proposed circuit is the Op-Amp, the frequency responses of the designed Op-Amp are shown in Figures 10 and 11.

As shown in those figures, the bandwidth of the Op-Amp is 346 MHz and has a phase margin of 65°. Also, the DC gain of this amplifier is about 78 dB. The output switching of this Op-Amp is from -0.86 to 0.86 volt, which will determine and limit the output changes in the proposed circuit. The values of the parameters used in the designed circuit are given in the Table 1.

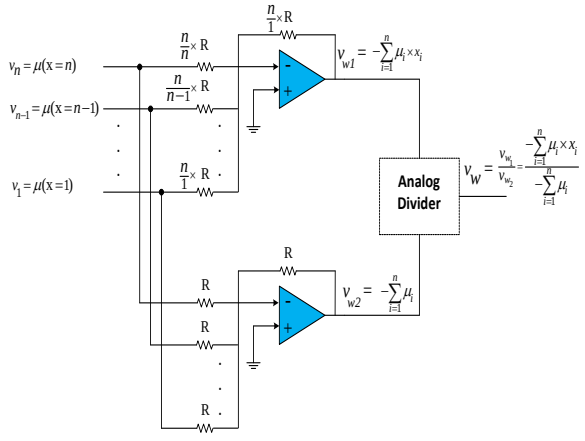


Figure 8. Hardware implementation of the WA method as the defuzzification procedure for $R_x = n$

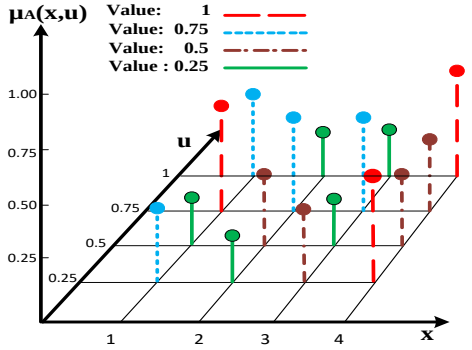


Figure 9. Example of an T2F for discrete universe of discourse

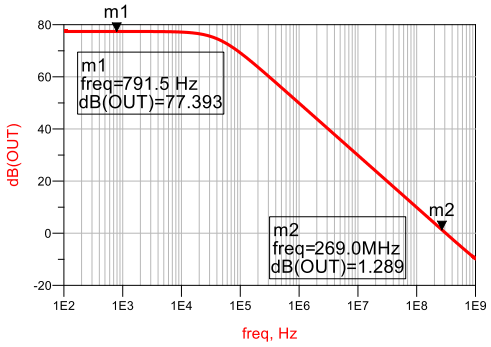


Figure 10. Frequency response of designed Op-Amp (amplitud)

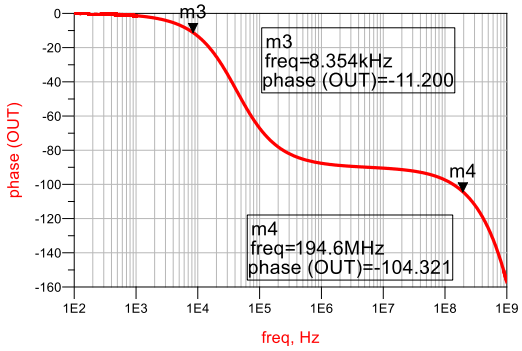


Figure 11. Frequency response of designed Op-Amp (Phase)

TABLE 1. The values of the parameters used in the designed circuit.

# Parameters	Parameter	Value
1	D	$10 * 10^{-9}\ m$
2	μ_v	$5 * 10^{-14}\ \frac{m^2}{VS}$
3	R_{on}	100Ω
4	R_{off}	$16\ K\Omega$
5	R_b	$10\ K\Omega$
6	R	$1\ K\Omega$
7	R_f	$8\ K\Omega$

To read the memristor state and encode it to voltage, the V_{Read} must be applied very quickly (similar to the impulse function) so that the memristor state does not change. The pulse width is about $1\mu s$, which can be easily produced by digital and analog systems.

The voltage of Op-Amp input (v_{m1}) is zero due to the virtual ground of the positive base of the Op-Amp and the negative feedback. In the first stage, the input voltage value is given to the Op-Amp according to Figure 12. The resulted Op-Amp output is shown in Figure 13.

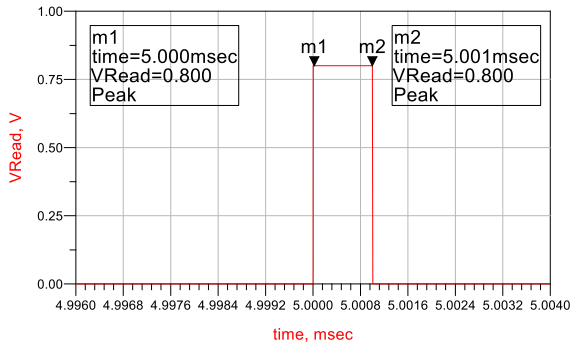


Figure 12. Reading voltage (V_{Read}) for training and pulse width

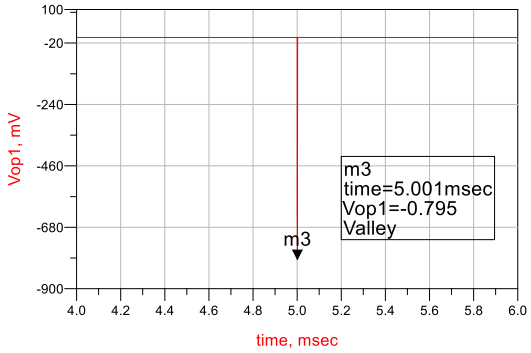


Figure 13. The output voltage propotional to the first stage of Op-Amp

The value of circuit parameters are designed in a way that the output of the second Op-Amp stage be zero, when no training is done. The voltage of V_{Bias} is equal to the input voltage V_{Read} and the values of the voltages are shown in the Figures 14 and 15.

After performing the training process and applying training pulses with the appropriate pulse width, the values stored in the memristor-crossbar structure is shown as Figure 16.

As shown in those figures, the outputs of Figures 14 and 15 are in accordance with Equations (7) and (8).

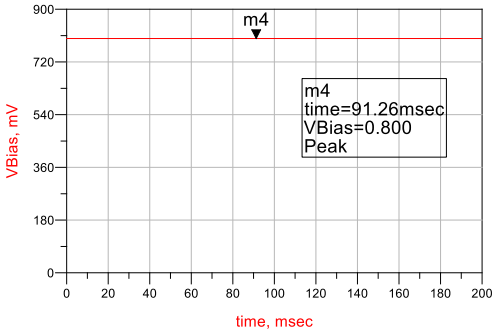


Figure 14. Bias node voltage waveform (V_{Bias})

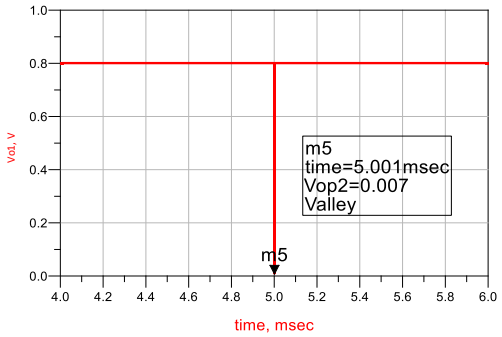


Figure 15. The output voltage propotional to the first stage of Op-Amp

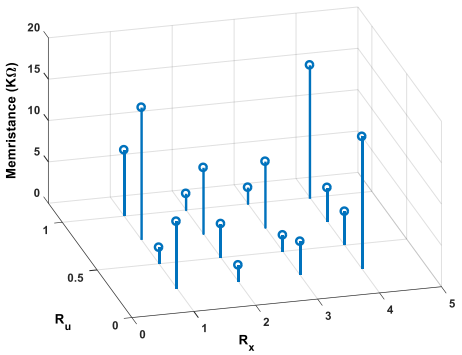


Figure 16. Memristance of memristors in memristor-crossbar

Also, the pattern in Figure 16 is similar to Figure 9. Therefore, it can be claimed that the implementation has been done with great accuracy.

Finally, it should be mentioned that providing the hardware that has learning ability is another challenge that Bi-Inspired systems are facing. The learning capability enables the hardware to be used for a variety of applications. Today, the parameters are mainly obtained by a host system, and then these parameters are implemented on the hardware. Therefore, the challenge of learning on independent hardware is considered in this paper. Moreover, in our proposed hardware, the changes in the memristance of memristors can be done at any time of system implementation while the membership degree values are programmed in the memristance values of memristors.

6. CONCLUSION

Due to the development and growing use of soft computing methods to solve complex engineering problems, it is necessary to provide hardware platforms souitable for the implementation of these algorithms. Memristor is a resistor with memory where memristor-crossbar structures are used to implement artificial neural networks, fuzzy systems, and neuro-fuzzy systems. Therefore, the goal of a suitable hardware platform has been achieved to some extent using nanoscale memristor-crossbar structures. In this paper, first, the T2F membership functions were investigated, and then the memristive hardware related to this type of fuzzy set was presented. Finally, it can be noted that, utilizing the proposed hardware structure, any number of membership functions with any shape and resolution can be implemented without the need for a host system. This feature become more important in adaptive systems. Future work includes the complete hardware implementation of a type-2 fuzzy computing system.

7. REFERENCES

1. Tanaka, K. and Wang, H.O., "Fuzzy control systems design and analysis: A linear matrix inequality approach, John Wiley & Sons, (2004).
2. Terano, T., Asai, K. and Sugeno, M., "Applied fuzzy systems, Academic Press, (2014).
3. Zadeh, L.A., Fu, K.-S. and Tanaka, K., "Fuzzy sets and their applications to cognitive and decision processes: Proceedings of the us–japan seminar on fuzzy sets and their applications, held at the university of california, berkeley, california, july 1-4, 1974, Academic press, (2014).
4. Srikrishna, A., Reddy, B.E. and Srinivas, V.S., Detection of lesion in mammogram images using differential evolution based automatic fuzzy clustering, in Computational intelligence techniques in health care. 2016, Springer.61-68. doi.org/10.1007/978-981-10-0308-0_5
5. Klidbary, S.H., Shouraki, S.B., Ghaffari, A. and Kourabbaslou, S.S., "Outlier robust fuzzy active learning method (alm)", in 2017 7th international conference on computer and knowledge engineering (ICCKE), IEEE. (2017), 347-352. doi: 10.1109/ICCKE.2017.8167903
6. Mendel, J.M., Uncertain rule-based fuzzy systems, in Introduction and new directions. 2017, Springer.684. doi.org/10.1007/978-3-319-51370-6
7. Mendel, J.M. and John, R.B., "Type-2 fuzzy sets made simple", ***IEEE Transactions on Fuzzy Systems***, Vol. 10, No. 2, (2002), 117-127, doi: 10.1109/91.995115.
8. John, R.I., Innocent, P.R. and Barnes, M., "Neuro-fuzzy clustering of radiographic tibia image data using type 2 fuzzy sets", ***Information Sciences***, Vol. 125, No. 1-4, (2000), 65-82, doi.org/10.1016/S0020-0255(00)00009-8.
9. Ozen, T. and Garibaldi, J.M., "Investigating adaptation in type-2 fuzzy logic systems applied to umbilical acid-base assessment", in Proceedings of 2003 European Symposium on Intelligent Technologies (EUNITE 2003). (2003), 289-294.
10. Figueroa, J., Posada, J., Soriano, J., Melgarejo, M. and Rojas, S., "A type-2 fuzzy controller for tracking mobile objects in the context of robotic soccer games", in The 14th IEEE International Conference on Fuzzy Systems, 2005. FUZZ'05., IEEE. (2005), 359-364. doi: 10.1109/FUZZY.2005.1452420
11. Melin, P., Mendoza, O. and Castillo, O., "An improved method for edge detection based on interval type-2 fuzzy logic", ***Expert Systems with Applications***, Vol. 37, No. 12, (2010), 8527-8535, doi.org/10.1016/j.eswa.2010.05.023.
12. Kim, D., " An implementation of fuzzy logic controller on the reconfigurable fpga system", ***IEEE Transactions on industrial Electronics***, Vol. 47, No. 3, (2000), 703-715, doi: 10.1109/41.847911.
13. Klidbary, S.H., Shouraki, S.B. and Linares-Barranco, B., "Digital hardware realization of a novel adaptive ink drop spread operator and its application in modeling and classification and on-chip training", ***International Journal of Machine Learning and Cybernetics***, Vol. 10, No. 9, (2019), 2541-2561, doi: doi.org/10.1007/s13042-018-0890-x.
14. Hung, D.L., "Dedicated digital fuzzy hardware", ***IEEE Micro***, Vol. 15, No. 4, (1995), 31-39, doi: 10.1109/40.400640.
15. Chua, L.O. and Kang, S.M., "Memristive devices and systems", ***Proceedings of the IEEE***, Vol. 64, No. 2, (1976), 209-223, doi: 10.1109/PROC.1976.10092..
16. Chua, L., "Memristor-the missing circuit element", ***IEEE Transactions on Circuit Theory***, Vol. 18, No. 5, (1971), 507-519, doi: 10.1109/TCT.1971.1083337.
17. Strukov, D.B., Snider, G.S., Stewart, D.R. and Williams, R.S., "The missing memristor found", ***Nature***, Vol. 453, No. 7191, (2008), 80-83, doi: doi.org/10.1038/nature06932.
18. Waser, R. and Aono, M., Nanoionics-based resistive switching memories, in Nanoscience and technology: A collection of reviews from nature journals. 2010, World Scientific.158-165.
19. Kuekes, P., "Material implication: Digital logic with memristors", in Memristor and memristive systems symposium. Vol. 21, (2008).
20. Tarkhan, M., Maymandi-Nejad, M., Haghzad Klidbary, S. and Bagheri Shouraki, S., "A bridge technique for memristor state programming", ***International Journal of Electronics***, Vol. 107, No. 6, (2020), 1015-1030, doi.org/10.1080/00207217.2019.1692371.
21. Mouttet, B., " Proposal for memristors in signal processing", in International Conference on Nano-Networks, Springer. Vol., No., (2008), 11-13. doi: 10.1007/978-3-642-02427-6_3
22. Merrikh-Bayat, F. and Shouraki, S.B., "Memristor-based circuits for performing basic arithmetic operations", ***Procedia Computer***

Science, Vol. 3, (2011), 128-132, doi.org/10.1016/j.procs.2010.12.022.

23. Amer, S., Madian, A.H. and Emara, A.S., " Memristor-based center-of-gravity (COG) defuzzifier circuit", in 2015 European Conference on Circuit Theory and Design (ECCTD), IEEE. (2015), 1-4. doi: 10.1109/ECCTD.2015.7300099
24. Amer, S., Madian, A., ElSayed, H. and Emara, A., "Effect of the memristor threshold current on memristor-based min-max circuits", in 2016 5th International Conference on Modern Circuits and Systems Technologies (MOCAST), IEEE. (2016), 1-4. doi: 10.1109/MOCAST.2016.7495104
25. Merrikh-Bayat, F., Shouraki, S.B. and Merrikh-Bayat, F., "Memristor crossbar-based hardware implementation of fuzzy membership functions", in 2011 eighth international conference on fuzzy systems and knowledge discovery (FSKD), IEEE. Vol. 1, (2011), 645-649.
26. Merrikh-Bayat, F. and Shouraki, S.B., "Memristive neuro-fuzzy system", ***IEEE Transactions on Cybernetics***, Vol. 43, No. 1, (2012), 269-285, doi.
27. Merrikh-Bayat, F., Shouraki, S.B. and Rohani, A., "Memristor crossbar-based hardware implementation of the ids method", ***IEEE Transactions on Fuzzy Systems***, Vol. 19, No. 6, (2011), 1083-1096, doi: 10.1109/TFUZZ.2011.2160024.
28. Klidbary, S.H. and Shouraki, S.B., "A novel adaptive learning algorithm for low-dimensional feature space using memristor-crossbar implementation and on-chip training", ***Applied Intelligence***, Vol. 48, No. 11, (2018), 4174-4191, doi.org/10.1007/s10489-018-1202-6.
29. Klidbary, S.H., Shouraki, S.B. and Afrakoti, I.E.P., "An adaptive efficient memristive ink drop spread (IDS) computing system", ***Neural Computing and Applications***, Vol. 31, No. 11, (2019), 7733-7754, doi.org/10.1007/s00521-018-3604-0.
30. Afrakoti, I.E.P., Shouraki, S.B., Bayat, F.M. and Gholami, M., "Using a memristor crossbar structure to implement a novel adaptive real-time fuzzy modeling algorithm", ***Fuzzy Sets and Systems***, Vol. 307, (2017), 115-128, doi.org/10.1016/j.fss.2016.02.016.
31. Adhikari, S.P., Yang, C., Kim, H. and Chua, L.O., "Memristor bridge synapse-based neural network and its learning", ***IEEE Transactions on Neural Networks and Learning Systems***, Vol. 23, No. 9, (2012), 1426-1435, doi: 10.1109/TNNLS.2012.2204770.
32. Hasan, R., Taha, T.M. and Yakopcic, C., "On-chip training of memristor crossbar based multi-layer neural networks", ***Microelectronics Journal***, Vol. 66, (2017), 31-40, doi.org/10.1016/j.mejo.2017.05.005.
33. Li, T., Duan, S., Liu, J., Wang, L. and Huang, T., "A spintronic memristor-based neural network with radial basis function for robotic manipulator control implementation", ***IEEE Transactions on Systems, Man, and Cybernetics: Systems***, Vol. 46, No. 4, (2015), 582-588, doi: 10.1109/TSMC.2015.2453138.
34. Alibart, F., Zamanidoost, E. and Strukov, D.B., "Pattern classification by memristive crossbar circuits using ex situ and in situ training", ***Nature Communications***, Vol. 4, No. 1, (2013), 1-7, doi: 10.1038/ncomms3072.
35. Karnik, N.N. and Mendel, J.M., "Introduction to type-2 fuzzy logic systems", in 1998 IEEE international conference on fuzzy systems proceedings. IEEE world congress on computational intelligence (Cat. No. 98CH36228), IEEE. Vol. 2, (1998), 915-920.
36. Karnik, N.N., Mendel, J.M. and Liang, Q., "Type-2 fuzzy logic systems", ***IEEE Transactions on Fuzzy Systems***, Vol. 7, No. 6, (1999), 643-658, doi: 10.1109/FUZZY.1998.686240.
37. Biolek, D., Di Ventra, M. and Pershin, Y.V., "Reliable spice simulations of memristors, memcapacitors and meminductors", ***arXiv preprint arXiv:1307.2717***, (2013).

Persian Abstract

چکیده

استفاده از تکنیک های فازی، به ویژه فازی نوع-۲، یکی از روش های بسیار پرکاربرد در یادگیری ماشین برای مدل سازی عدم قطعیت است. علاوه بر ارائه الگوریتم، قابلیت پیاده سازی سخت افزاری و همچنین عملکرد مناسب در کاربردهای زمان واقعی از چالش های مهم هستند. استفاده از بسترهای سخت افزاری که شباهت های بیولوژیکی دارند و از نظر حجم پیاده سازی با سیستم های عصبی انسان قابل مقایسه هستند همیشه مورد توجه بوده است. ممریستور یکی المان های نوظهور برای پیاده سازی الگوریتم های مبتنی بر منطق فازی است. در این المان، با اعمال جریان و انتخاب جهت مناسب برای جریان اعمال شده ، مقاومت ممریستور (ممریستنس) افزایش یا کاهش می یابد. پیاده سازی های مختلفی از سیستم های فازی نوع-۱ وجود دارد، اما هیچ پیاده سازی از سیستم های فازی نوع-۲ مبتنی بر مموریستور انجام نشده است. در این مقاله، از ساختارهای متقاطع ممریستور برای پیاده سازی توابع عضویت فازی نوع-۲ استفاده شده است. در سخت افزار پیشنهادی، توابع عضویت می توانند از هر شکل و رزولوشنی برخوردار باشند. سخت افزاری پیشنهادی که برای توابع عضویت فازی نوع-۲ ارائه شده است، قابلیت یادگیری دارد (قابلیت یادگیری روی تراشه بدون نیازی به سیستم میزبان). علاوه بر این، سخت افزار پیشنهادی آنالوگ بوده و می تواند به عنوان مبنایی در ساخت سیستم های تکاملی مورد استفاده قرار گیرد. ضمناً، برای بررسی عملکرد سخت افزار پیشنهادی، شبیه سازی هایی مبتنی بر ممریستورانجام داده شده است.

International Journal of Engineering

Journal Homepage: www.ije.ir

Effects of Cold Rolling and Annealing Time on Fatigue Resistance of AA5052 Aluminum Alloy

Kusmono*, C. Bora, U. A. Salim

Department of Mechanical and Industrial Engineering, Faculty of Engineering, Universitas Gadjah Mada, Indonesia

P A P E R I N F O

Paper history:
Received 16 May 2021
Received in revised form 29 July 2021
Accepted 02 August 2021

Keywords:
AA5052 Aluminum Alloy
Cold Rolling
Annealing
Fatigue Resistance

A B S T R A C T

In the present study, the influences of cold rolling and annealing time on the fatigue crack propagation behavior of AA5052 aluminum alloy were investigated. The alloy sheet was cold-rolled under different rolling reductions, i.e., 0, 15, 30, and 45%. A 45% as-rolled specimen was then annealed at 370°C under different annealing times, i.e., for 2, 4, and 6 h. The microstructure evolutions after the cold rolling and annealing treatments were also examined using optical microscopy whereas the fatigue crack propagation behavior was characterized by using a fatigue test. Results showed that severely elongated grains were observed with increasing the rolling reduction. The elongated microstructures were changed into equaxial structures due to recrystallization during annealing treatment. The fatigue life was decreased drastically by increasing the rolling reduction but increased significantly with increasing annealing time. The fatigue life of the alloy was reduced by 93% when cold-rolled up to 45%. On the other hand, the fatigue life of the 45% rolled samples was increased significantly by 412% when annealed at 370ºC for 6 h.

***Doi**: 10.5829/ije.2021.34.09c.16*

1. INTRODUCTION[1]

Due to its excellent properties including lightweight, processability, physical characteristics, corrosion resistance, and low cost, aluminum-magnesium alloys have been widely applied in the aerospace, automotive, shipbuilding, and construction industries [1,2]. One of the most widely used Al-Mg alloys is the AA5052 aluminum alloy. However, this alloy is non-heat treatable in that its strength is not high enough to be used for structural applications [3]. To improve its mechanical properties, there are two well-known methods, namely micro-alloying and strain hardening [4]. It is reported that the mechanical characteristics of aluminum alloys can be improved by the application of severe plastic deformations (SPD) such as accumulative roll-bonding, cryogenic rolling, and equal channel angular pressing [5,6]. Cold rolling has been proven to be one of the most widely used methods of strain hardening for Al-Mg aluminum alloys due to its simplicity, ease of processing, and low cost. Cold rolling can significantly enhance the strength and hardness of Al-Mg aluminum alloy because of the improved dislocation density [7]. However, the ductility of the cold-rolled aluminum alloys is lower which limits its application due to poor forming ability. Tabatabaeia and Zarasvand [8] studied the influences of cold bulge forming speed on thickness reduction and tensile properties of aluminum alloys. Their results showed that the tensile and yield strength of aluminum alloys were increased after cold bulge forming. Moreover, a more uniform thickness was achieved by reducing the forming speed. To enhance the ductility, an annealing treatment is often applied for cold-rolled Al-Mg aluminum alloys [9]. In our previous studies [10], it was reported that the strength and hardness of AA5052 aluminum alloy were significantly improved with rolling reduction but the ductility decreased. Furthermore, the annealing process decreased both strength and hardness but increased ductility. Several factors play a role in the success of annealings such as temperature, holding time, and heating rate of the process.

*Corresponding Author Institutional Email: *kusmono@ugm.ac.id* (Kusmono)

Please cite this article as: Kusmono, C. Bora, U. A. Salim, Effects of Cold Rolling and Annealing Time on Fatigue Resistance of AA5052 Aluminum Alloy, International Journal of Engineering, Transactions C: Aspects, Vol. 34, No. 09, (2021) 2189-2197

Manufacturing components made of AA5052 which are used in several applications such as aerospace, automotive, and shipbuilding are often subjected to cyclic loading which can result in fatigue failure. In connection with this, several researchers have conducted many studies on the fatigue behavior of AA5052 aluminum alloy. Azadi et al. [11] investigated the effects of various parameters (maximum temperature, dwell time, thermo-mechanical loading factor) on the out-of-phase thermo-mechanical fatigue lifetime of the A356.0 casting aluminum alloy. They found that the maximum temperature and thermo-mechanical loading factor had a more significant effect on the lifetime compared to the dwell time. Li et al. [4] studied the effect of annealing treatment on the microstructure, tensile strength, and fatigue crack propagation behavior of Al-Mg-Sc alloy sheets. The sheets were hot-rolled at 470ºC with the 77% rolling reduction and then cold-rolled with 65% reduction. The sheets were then annealed at 340 and 500ºC for 1 h. They found that the microstructure, tensile strength, and fatigue crack growth rate were greatly dependent on the annealing temperature. Studies on the fatigue resistance of 5083-O aluminum alloy cold-rolled in parallel (PD) and in vertical to rolling direction (VD) were done by Ma et al. [12]. It was reported that the fatigue resistance of PD specimens was lower than that of the VD specimen due to the primary crack propagation throughout the chain-shaped coarse inclusions in PD specimens. Mughrabi and Hoppel [13] demonstrated that the fatigue properties of metal alloys are strongly affected by physical metallurgical properties such as grain size. Grain size affects remarkably both fatigue-crack propagation characteristics of alloys. Mughrabi and Hoppel [13] and Kim and Bidwell [14] reported that a microstructure with a relatively larger grain size has a higher fatigue-crack-propagation threshold (ΔK_{th}) and lower crack-growth rate. Shou et al. [15] investigated the effect of grain size on the fatigue crack propagation behavior of the 2524-T3 aluminum alloy. They found that the increased fatigue resistance of alloys with larger grain sizes is due to crack closure induced by both surface roughness and plasticity. Research works on the fatigue characteristics of aluminum alloys have been studied extensively, but published data on the influences of cold rolling and annealing treatment of AA5052 aluminum alloy have been limited.

In this study, the AA5052 alloy sheet was cold-rolled with different rolling reductions, namely 15, 30, and 45%. The sample with 45% reduction was then annealed at 370°C for different holding times, namely 2, 4, and 6 h. This work, the influences of cold rolling and annealing time on the fatigue crack propagation characteristic of AA5052 aluminum alloy were discussed. The fatigue fracture surfaces of these alloys were also characterized.

2. MATERIAL AND EXPERIMENTAL

The material used in this study was an AA5052 plate with its chemical composition is given in Table 1. In this investigation, the plate with 3 mm in thickness was firstly homogenized at 470°C for 6 h with a heating rate of 5°C/min and then cooled in air. The homogenized sheet was then cold-rolled at room temperature under different rolling reductions, namely 0, 15, 30, and 45%. The cold-rolled sheet with 45% reduction was then annealed at 370°C for different times, i.e., 2, 4 and 6 h with a heating rate of 5°C/min. The schematic illustration of the cold rolling and annealing processes was presented in Figure 1.

The microstructures of as-received, cold-rolled, and annealed specimens were observed using optical microscopy in a short transverse direction. The analysis of phases was carried out using an X-ray diffractometer (X'pert MPD) with Cu Kα X-ray source and 2θ range from 20 to 90º at a scanning rate of 0.03°/s. The fatigue crack growth test was conducted by using a servo-hydraulic universal testing machine with the center-crack tension (CCT) specimens taken from the sheet in the rolling direction as shown in Figure 2. Electric discharge machining was used to prepare the notches and the surface of the sample under test was polished. All tests were carried out at a stress ratio of R = 0.1, and with a sinusoidal loading frequency of 11 Hz at room temperature. A 20% of yield stress was selected as stress level and the fatigue crack growth rate (da/dN) of the Paris power law was determined following the Secant method below:

$$\left(\frac{da}{dN}\right)_{\overline{a}} = \frac{a_{i+1}-a_i}{N_{i+1}-N_i} \quad (1)$$

$$\overline{a} = \left(\frac{a_{i+1}-a_i}{2}\right) \quad (2)$$

where, $\overline{a}$ is average crack length, subscripts i and $(i+1)$ represent i_{th} and $(i+1)_{\text{th}}$ cycle. The stress intensity factor range, ΔK for center crack tension (CCT) geometry was determined as follows:

$$\Delta K = \frac{\Delta P}{B}\left[\sqrt{\frac{\pi\alpha}{2W}}\, sec\frac{\pi\alpha}{2}\right] \quad (3)$$

$$\Delta P = P_{max} - P_{min} \quad (4)$$

where P is the load, B and W are the thickness and width

TABLE 1. Chemical compositions of AA5052 aluminum alloy

Chemical composition (wt.%)					
Mg	**Fe**	**Cr**	**Si**	**Mn**	**Al**
2.37	0.31	0.22	0.11	0.08	Bal.

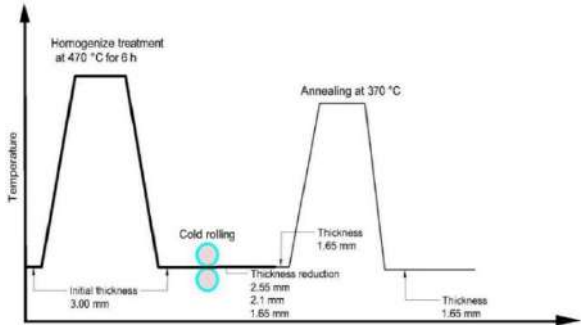


Figure 1. Schematic illustration of both cold rolling and annealing processes

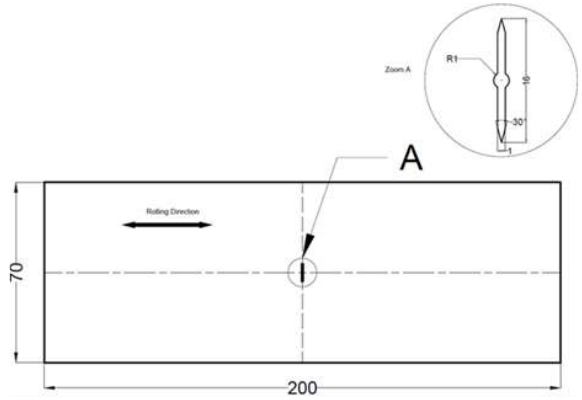


Figure 2. Centre-crack tension (CCT) specimen

of the sample, respectively; $\alpha = 2a/W$; and a is the length of half of the crack. The fatigue fracture surface of the specimens was examined using scanning electron microscopy (Carl Zeiss Evo Ma 10 SEM).

3. RESULTS AND DISCUSSION

3. 1. Microstructure Observations Figure 3a shows three principal directions in the alloy during cold rolling. Figures 3b, 3c, 3d, 3e, and 3f display the optical micrographs in the short transverse direction of as-received, as-rolled with 15, 30, 45% reduction, and as-rolled with 45% reduction followed by annealing for 6 h, respectively. The structure of the α(Al) matrix and many second phases dispersed within the grain are observed in all samples. Some equiaxed grains and elongated ones can be observed in Figure 3b. However, a larger elongated grain as a result of cold rolling is exhibited clearly in Figures 3c, 3d, and 3e. This indicates that the cold rolling led to a smaller grain size of alloy. An increase in the number of nucleation sites when higher rolling reductions are applied may be responsible for the smaller grain size [16]. Some equiaxed grains can be observed in the 45% as-rolled sample followed by annealing for 6 h in Figure 3f. This suggests that the

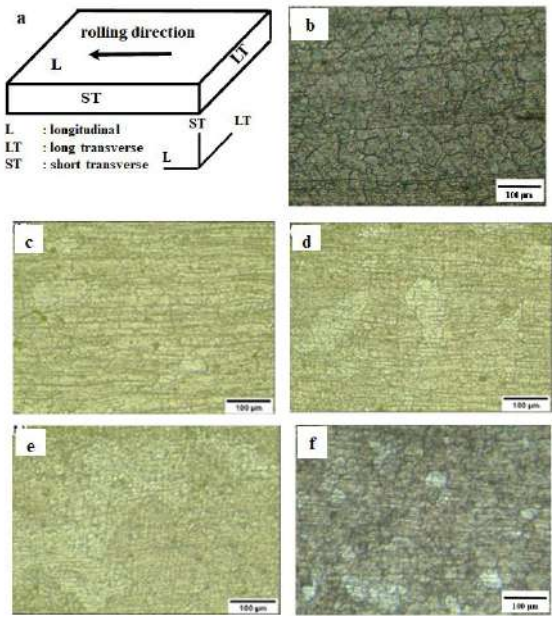


Figure 3. (a) Three principal directions in the alloy sheet with corresponding microstructure in short transverse: (b) as-received sample, (c) 15% as-rolled sample, (d) 30% as-rolled sample, (e) 45% as-rolled sample, (f) 45% as-rolled sample followed by annealing for 6 h

annealing for 6 h has changed from the elongated grains to the equiaxed ones. This may be related to the recrystallization phenomenon that occurred during annealing treatment [9].

3. 2. X-Ray Diffraction Analysis The XRD patterns of as-received and as-rolled samples with different thickness reductions are demonstrated in Figure 4. As shown in Figure 4, it can be observed that both as-received and as-rolled samples consist of similar phases, namely α(Al) matrix, and Al_3Fe. No other phases were observed. This suggests that the cold rolling does not change the phases in the AA5052 aluminum alloy. However, the intensity of diffraction peaks is increased with rolling reduction, especially for 45% reduction. This indicates that the amount of secondary phase (Al_3Fe) increases with increasing rolling reduction. An increase in cold rolling causes greater energy stored in the cold-rolled sample as indicated by an increase in the intensity of Al_3Fe as shown in the XRD results. In addition, cold rolling produces an increase in dislocation density. Similar observations were also demonstrated by previous researchers [9] where the cold rolling did not affect the microstructure but only increased the amount of the second phase as indicated by the enhanced intensity of peaks after cold rolling. Figure 5 displays the XRD patterns of as-rolled with 45% reduction and annealed samples for different holding times. It can be seen from

Figure 5 that all samples are composed of α (Al matrix and Al_3Fe as shown in Figure 4. This implies that the annealing treatment at 370°C for 2, 4, and 6 h does not affect the phase changes of as-rolled AA 5052 aluminum alloy. The annealing time does not influence the presence of Al_3Fe because the annealing treatment reduces the dislocation density due to the recrystallization process during annealing treatment.

3. 3. Fatigue Crack Propagation Behavior

Results of fatigue tests of the as-received and as-rolled alloy under different rolling reductions are displayed in Figure 6. It was found that the fatigue life of as-received alloy was around 1,050,000 cycles. A remarkable reduction in fatigue life is exhibited by a 15% cold-rolled sample where the number of cycles decreased to around 560,000 cycles. More reduction in fatigue life is also found for as-rolled samples with the rolling reductions of 30 and 45%. The number of cycles of as-rolled specimens

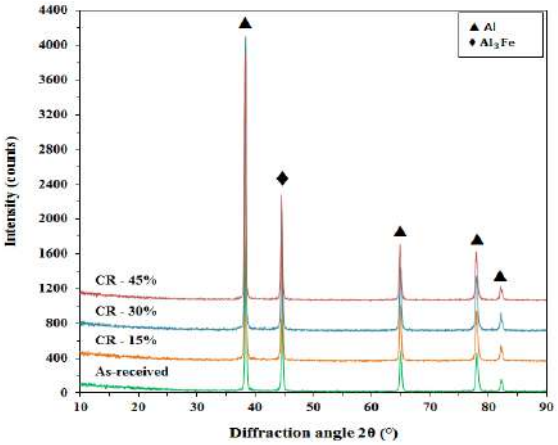


Figure 4. XRD patterns of as-received and as-rolled samples with different thickness reductions

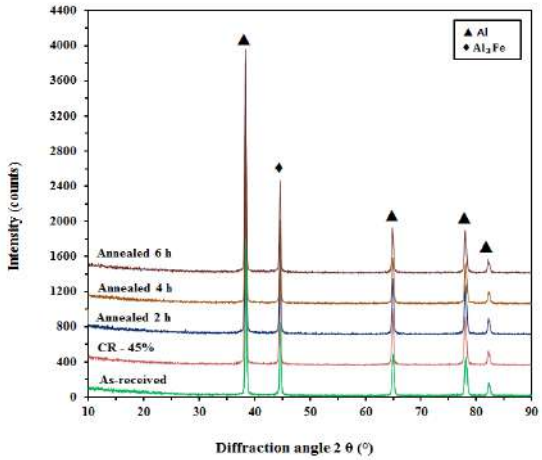


Figure 5. XRD patterns of as-rolled samples with 45% reduction subjected to annealing at different times

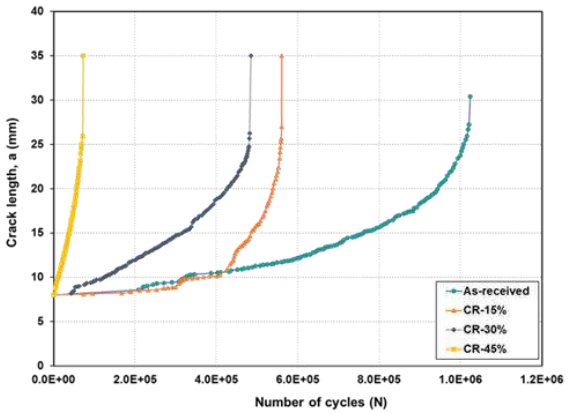


Figure 6. Fatigue life of as-received and as-rolled samples with different rolling reductions

with 30 and 45% reduction is 485,000 and 73,000 cycles, respectively. It can be concluded that the fatigue life of alloy is drastically reduced with rolling reduction. This is probably attributed to the smaller grain size and high tensile residual stress caused by cold rolling [17, 18]. These increases the local stress ratio and the effective stress intensity factor range and finally result in reduced fatigue life [18].

Figure 7 displays the fatigue crack growth rate as a function of the stress intensity factor range (ΔK) of as-received and as-rolled alloy with different rolling reductions. It is observed that the fatigue crack propagation curves reveal three different regions, i.e., I, II, and III. For both samples, region II shows a stable rate of fatigue crack growth following the Paris law as given by:

$$\frac{da}{dN} = C\ (\Delta K)^n \qquad (5)$$

where da/dN is the growth rate of fatigue cracks, and ΔK is the difference between the maximum and minimum stress intensity factors, (K_{max} – K_{min}) during the load cycle. By taking the trend line in region II, the values of C and n known as the Paris constant can be determined and the results are presented in Figure 8 and summarized in Table 2. C is the value of da/dN at ΔK = 1 MPa.m$^{0.5}$ and n represents the slope of the lines. The n value of aluminum alloy increased with rolling reduction. This indicates that the cold rolling increased the fatigue crack growth rate. The reduced fatigue performance is ascribed to the smaller grain size and tensile residual stress induced by cold rolling. These findings were consistent with previous results reported that the rate of fatigue crack growth of fine-grained commercially pure titanium induced by the cold rolling was higher compared to that of the coarse-grained samples due to the difference in the magnitude of crack closure of the samples [19-21]. The rate of fatigue crack growth of aluminum alloys is greatly influenced by grain size [15,22]. It has been reported by previous researchers that the higher rate of fatigue crack

propagation was indicated by a smaller grain size compared to alloys with a larger grain size [23]. This is related to the combined effects of grain boundaries, crack deflection, and crack closure due to crack surface roughness and plasticity [15].

Figure 9 demonstrates the fatigue life of 45% as-rolled alloy with different annealing holding times. It is found that the fatigue life for samples annealed for 2, 4, and 6 h are 213,000, 232,000, and 374,000 cycles, respectively. Compared to the un-annealed sample, there is a significant increase in fatigue life by 192, 218, and 412% for samples annealed for 2, 4, and 6 h, respectively. This indicates that the annealing treatment significantly enhances the fatigue life of an as-rolled alloy. This may be due to the more ductile nature of the annealed alloy compared to the rolled sample allowing a greater number of crack extensions before failure [24]. In addition, these are also attributed to both the reduced dislocation density and tensile residual stress because of recrystallization during annealing [25].

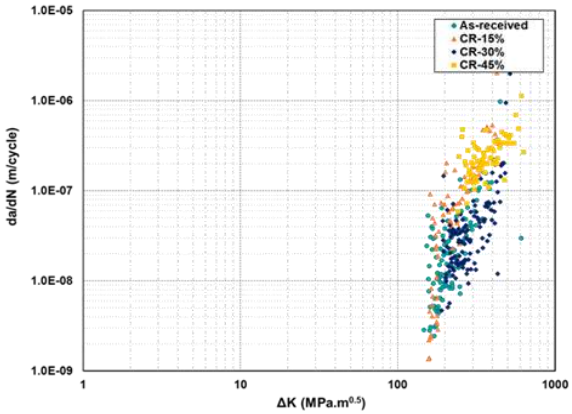


Figure 7. Fatigue crack propagation rate of as-received and as-rolled samples with different rolling reductions

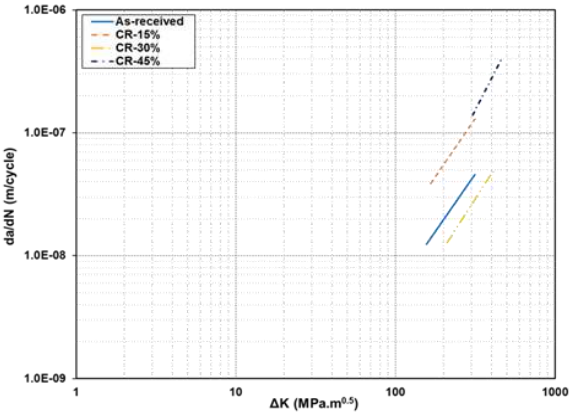


Figure 8. Trend line taken from region II of Figure 7

TABLE 2. Paris Constants of as-received, as-rolled samples and followed by annealing under different times

Materials	C	n
As-received	9.6351E-13	1.8726
CR-15%	2.3555E-12	1.8979
CR-30%	2.4204E-13	2.0317
CR-45%	6.9064E-14	2.5379
CR-45% + annealed for 2 h	5.7237E-14	2.3969
CR-45% + annealed for 4 h	4.3521E-13	2.2841
CR-45% + annealed for 6 h	3.5767E-13	2.0028

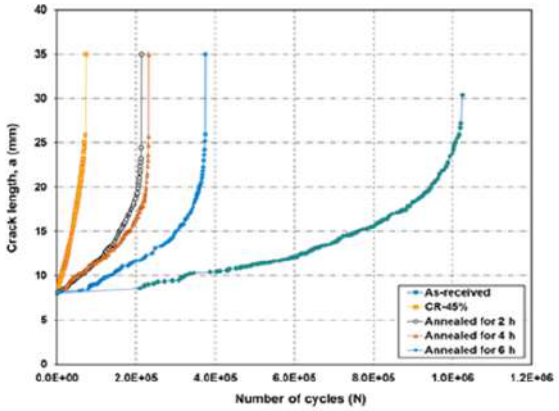


Figure 9. Fatigue life of as-received and as-rolled samples followed by annealing with different annealing times

Figure 10 demonstrates the fatigue crack growth rate as a function of stress intensity factor range (ΔK) of as-received and 45% as-rolled alloy for different annealing times. It can be seen that all samples also reveal the fatigue crack growth rate curves with three distinct regions, i.e., the region I, II, and III. All samples also exhibit the stable fatigue crack growth rate in region II and follow Paris law. The curve of da/dN versus ΔK taken from region II is presented in Figure 11 and Paris constants are determined that summarized in Table 2.

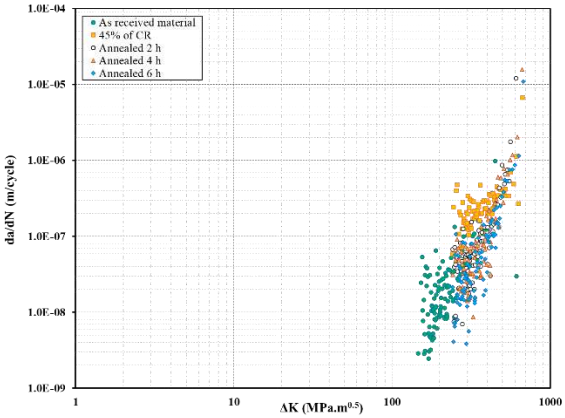


Figure 10. Fatigue crack propagation rate of as-received and 45% as-rolled samples with different annealing times

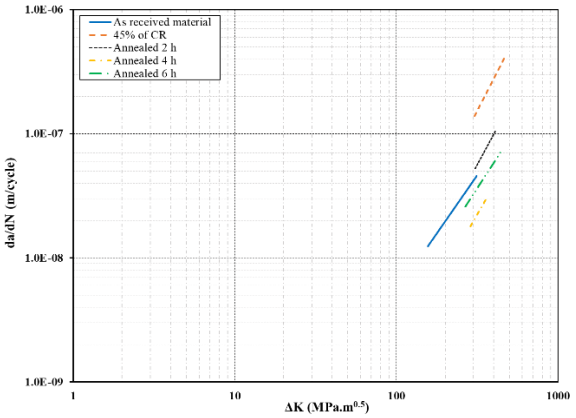


Figure 11. Trend line taken from region II of Figure 10

From Table 2, it can also be found that the *n* value is decreased with annealing time. This implies that the annealing treatment reduces the fatigue crack growth rate. This is associated with grain coarsening as a result of recrystallization during annealing. Furthermore, the grain coarsening can increase the intensity factor of the crack closure and decrease the driving force of change in the crack path. It is ascribed to the microstructure and possible contact between facets of rough cracks resulting in a low crack growth rate as indicated by lower *n* values after annealing treatment [26,27]. This may also be associated with the higher release rate of tensile residual stress with increasing annealing time producing a lower fatigue crack growth rate. In addition, The enhanced fatigue crack growth resistances for annealed alloys may be attributed to their high C values (but lower values of n) suggesting that the fatigue crack growth retardation occurs at high ΔK [28].

3. 4. Fatigue Fracture Surface Analysis Figure 12 shows the fatigue fractured surfaces of AA5052 aluminum alloy for as-received and as-rolled samples with 15 and 45% of rolling reduction. The crack propagates from the left to the right as shown in Figure 12 with the initial crack from the notch. From Figure 12, it can be seen that the fracture surface consisted of three different regions, namely an area with a flat surface that is characteristic of a fatigue fracture, an area with a rougher surface called a shear fracture or shear mode, and an area with a reduction in the area which is a feature of the final fracture. The fatigue surface area reflects the fatigue life of the fatigue-tested alloy [29]. Therefore, the larger the fatigue surface area confirms the higher the fatigue life. From Figure 12, it can be observed that the largest area of the fatigue fracture is observed for the as-received alloy followed by the cold-rolled alloy with 15% rolling reduction, and alloy with 45% rolling reduction. This confirms that the fatigue life of the alloy is reduced with an increase of rolling reduction as presented in Figure 6.

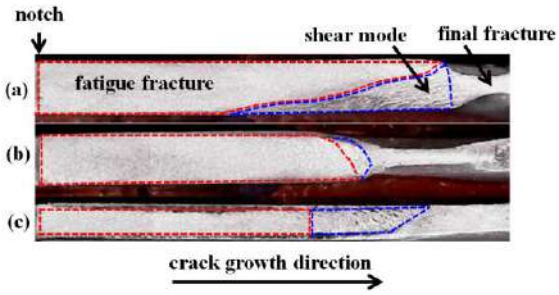


Figure 12. Fractured surfaces of AA 5052 aluminum alloy: (a) As-received, (b) CR-15%, (c). CR-45%

Figure 13 presents the fatigue fractured surfaces of AA 5052 aluminum alloys subjected to cold rolling with 45% and followed by annealing at different times of 2, 4, and 6 h. A similar to Figure 12, three different regions, namely the fatigue, shear, and final fractures can be observed. The larger area of the fatigue fracture is observed for the alloy annealed 6 h followed by 4 and 2 h. This suggests that the fatigue life of the alloy is increased with increasing annealing time. These results are consistent with the fatigue life as shown in Figure 9 before. From Figures 12 and 13, it can also be observed that the area of fatigue surface of the as-received alloy is still higher compared to that of all annealed alloys. This reflects that the fatigue life of as-received alloy is higher than that of cold-rolled alloy followed by annealing treatment at different times. In other words, the annealing treatment has not been able to restore the fatigue life of as-received alloy.

Figure 14 displays scanning electron micrographs of the fatigue fracture surfaces taken from stable crack propagation region of as-received, 45% as-rolled, and 45% as-rolled alloy annealed for 6 h. It can be seen that the presence of fatigue striation can be observed in the fatigue crack fractures of as-received and 45% as-rolled samples annealed for 6 h (Figures 14a and 14c). The smaller striation width can be seen in the fracture surface of the as-received sample compared to the 45% as-rolled alloy followed by annealing for 6 h. This supports the lower fatigue crack growth rate exhibited by the

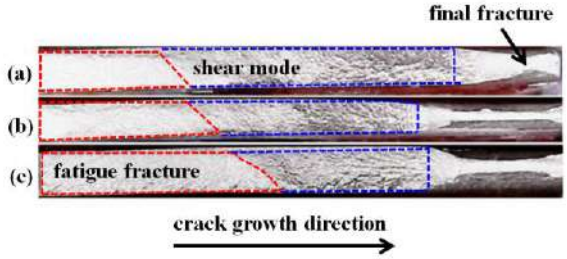


Figure 13. Fractured surfaces of AA 5052 aluminum alloy subjected to 45% of cold rolling and annealed at different times: (a) 2 h, (b) 4 h, (c) 6 h

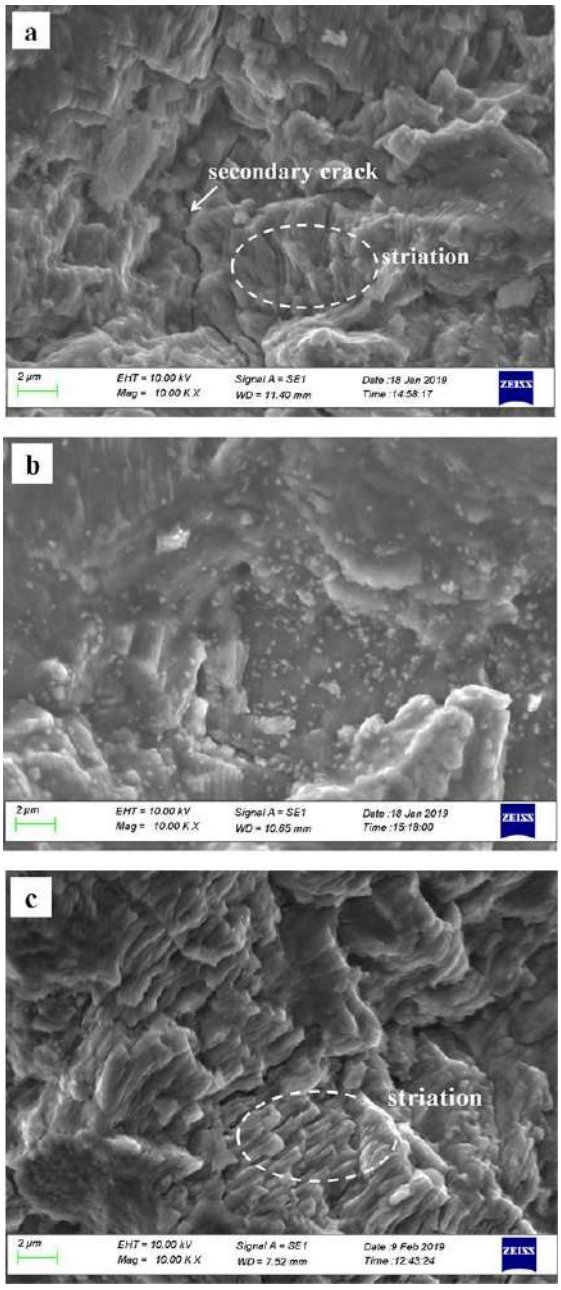


Figure 14. SEM images of fatigue fracture surfaces of (a) as-received alloy, (b) 45% as-rolled alloy, (c) 45% as-rolled alloy and annealed for 6 h

as-received sample compared to the 45% as-rolled sample annealed for 6 h, as shown in Figure 8 and Table 2. According to Matos et al. [30], striation width is the fatigue crack growth distance under each cyclic load reflecting the fatigue crack growth rate properly. In general, the smaller striation width corresponds to the lower fatigue crack growth rate [31,32]. Another feature seen in Figures 14a and 14c is the presence of the secondary crack obtained in the fracture surface of as-received and 45% as-rolled followed by annealing for 6 h. It is interesting to note from Figure 14b that poorly defined striation is revealed in the fatigue surface of 45% as-rolled AA5052 aluminum alloy. This indicates the transgranular cleavage fracture as typical of brittle fracture. This confirms the high fatigue crack propagation rate of 45% as-rolled AA 5052 aluminum alloy as shown in Figure 8 and Table 2.

4. CONCLUSIONS

In this study, the effects of cold rolling and annealing time on the microstructure and fatigue behavior of the AA5052 aluminum alloy was investigated and the conclusions were as follows:

- The severely elongated grains are observed with increasing rolling reduction. The elongated microstructures are changed to equaxial structures due to recrystallization during annealing treatment especially for an annealing holding time of 6 h.
- Both cold rolling and annealing processes do not change the phases in the AA5052 aluminum alloy.
- Both cold rolling and annealing time strongly influence the fatigue resistance of AA5052 aluminum alloy. The fatigue resistance decreases drastically as the rolling reduction increases. On the other hand, the fatigue resistance of cold-rolled AA5052 aluminum alloy increases significantly with increasing annealing time.

5. ACKNOWLEDGEMENTS

The financial support from Universitas Gadjah Mada, Yogyakarta, Indonesia through the research grant of Rekognisi Tugas Akhir (RTA) Program 2019 (contract no. 2129/UN1/DITLIT/DIT-LIT/LT/2019) is greatly appreciated.

6. REFERENCES

1. Xia, S.L., Ma, M., Zhang, J.X., Wang, W.X., Liu W.C. "Effect of heating rate on the microstructure, texture and tensile properties of continuous cast AA5083 aluminum alloy." ***Materials Science and Engineering A***, Vol. 609, (2014), 168-176, https://doi.org/10.1016/j.msea.2014.05.002
2. Panagopoulos, C.N., Georgiou, E.P. "Cold rolling and lubricated wear of 5083 aluminium alloy." ***Materials and Design***, Vol. 31, (2010), 1050-1055, https://doi.org/10.1016/j.matdes.2009.09.056.
3. Lin, S., Nie, Z., Huang, H., Li, B. "Annealing behavior of a modified 5083 aluminum alloy." ***Materials and Design,*** Vol. 31, (2010), 1607-1612, https://doi.org/10.1016/j.matdes.2009.09.004.
4. Li, M., Pan, Q., Shi, Y., Wang, Y. "Microstructure dependent fatigue crack growth in Al-Mg-Sc alloy." ***Materials Science & Engineering A***, Vol. 611, (2014), 142-151, https://doi.org/10.1016/j.msea.2014.05.087.

5. Liu, J.T., Morris J. "Recrystallization microstructures and textures in AA 5052 continuous cast and direct chill cast aluminum alloy." ***Materials Science and Engineering A***, Vol. 385, (2004), 342-351, https://doi.org/10.1016/j.msea.2004.06.070.
6. Daryadel, M. "Study on equal channel angular pressing process of AA7075 with copper casing by finite element-response Surface Couple Method." ***International Journal of Engineering: Transactions C: Aspects***, Vol. 33, No. 12, (2020) 2538-2548, DOI:10.5829/ije.2020.33.12c.15.
7. Morrisa, D.G., Muñoz-Morrisa, M.A. "Microstructure of severely deformed Al-3Mg and its evolution during annealing." ***Acta Materialia***, Vol. 50, No. 16, (2002), 4047-4060, https://doi.org/10.1016/S1359-6454(02)00203-3.
8. Tabatabaeia, S.M.R., Zarasvand, K.A. "Investigating the effects of cold bulge forming speed on thickness variation and mechanical properties of aluminum alloys: Experimental and Numerical." ***International Journal of Engineering, Transaction C: Aspects***, Vol. 31, No. 9, (2018), 1602-1608, DOI: 10.5829/ije.2018.31.09c.17.
9. Wang, B., Chen, X., Pan, F., Mao, J., Fang, Y. "Effects of cold rolling and heat treatment on microstructure and mechanical properties of AA 5052 aluminum alloy." ***Transaction Nonferrous Metal Society of China*** Vol. 25, (2015), 2481-2489, https://doi.org/10.1016/S1003-6326(15)63866-3.
10. Kusmono, Bora, C., Salim, U.A. "Effects of cold rolling (CR) and annealing time on microstructure and mechanical properties of AA 5052 aluminum alloy." Metallurgy, Vol. 59, No. 4, (2020), 485-488.
11. Azadi, M., Farrahi, G.H., Winter, G., Eichlseder, W. "The effect of various parameters on out-of-phase thermo-mechanical fatigue lifetime of A356.0 cast aluminum alloy." ***International Journal of Engineering: Transactions C: Aspects,*** Vol. 26, No. 12, (2013), 1461-1470, DOI: 10.5829/idosi.ije.2013.26.12c.06.
12. Ma, M., Zhang, J., Yi, D., Wang, B. "Investigation of high-cycle fatigue and fatigue crack propagation characteristic in 5083-O aluminum alloy." ***International Journal of Fatigue***, Vol. 126, (2019), 357-368, https://doi.org/10.1016/j.ijfatigue.2019.05.020.
13. Mughrabi H, Hoppel HW. Cyclic deformation and fatigue properties of ultrafine grain size materials: current status and some criteria for improvement of the fatigue resistance. ***Material Research Soceity Symposium Proceedings,*** (2001), 634. B2.1.1–B2.1.12.
14. Kim Y-W, Bidwell LR. Effects of microstructure and aging treatment on the fatigue crack growth behaviour of high strength P/M aluminum alloy X7091. In: High-strength powder metallurgy aluminum alloys: proceedings of a symposium sponsored by the Powder Metallurgy Committee of the Metallurgical Society of AIME, held at the 111th AIME Annual Meeting, Dallas, Texas, February 17–18, 1982/edited by Michael J. Koczak, Gregory J. Hildeman; 107-124.
15. Shou, W.B., Yi, D.Q., Liu, H.Q., Tang, C., Shen, F.H., Wang, B. "Effect of grain size on the fatigue crack growth behavior of 2524- T3 aluminum alloy." ***Archive of Civil and Mechanical Engineering***, Vol. 16, (2016), 304-312, https://doi.org/10.1016/j.acme.2016.01.004.
16. Nah, J.J., Kang, H.G., Huh, M.Y., Engler, O." Effect of strain states during cold rolling on the recrystallized grain size in an aluminum alloy." ***Scripta Materialia*** Vol. 58, (2008), 500-503, https://doi.org/10.1016/j.scriptamat.2007.10.049.
17. Liu, S.Z., Minakawa, K., Scholtes, B., Mcevily, A.J. "The effect of cold rolling on the fatigue properties of Ti-6Al-4V." ***Metallurgical Transactions A,*** Vol. 16, (1985), 144-145, https://doi.org/10.1007/BF02656725.
18. Korsgren, P., Sperle, J.O., Trogen, H. "Influence of shearing and punching on the fatigue strength of hot rolled steel sheet." ***Scandinavian Journal of Metallurgy***, Vol. 18, (1989), 203-210.
19. Kikuchi, S., Mori, T., Kubozono, H., Nakai, Y., Kawabata, M.O., Ameyama, K. "Evaluation of near-threshold fatigue crack propagation in harmonic-structured CP titanium with a bimodal grain size distribution. ***"Engineering Fracture Mechanics***, Vol.181, (2017), 77-86, https://doi.org/10.1016/j.engfracmech.2017.06.026.
20. Kikuchi, S., Ueno, A., Akebono, H. "Combined effects of low temperature nitriding and cold rolling on fatigue properties of commercially pure titanium." ***International Journal of Fatigue***, Vol. 139, (2020), 105772, https://doi.org/10.1016/j.ijfatigue.2020.105772.
21. Nakai, Y., Kikuchi, S., Osaki, K., Kawabata, M.O., Ameyama, K. "Effects of rolling reduction and direction on fatigue crack propagation in commercially pure titanium with harmonic structure." ***International Journal of Fatigue***, Vol. 143, (2021), 106018, https://doi.org/10.1016/j.ijfatigue.2020.106018.
22. Yin, D., Liu, H., Chen, Y., Yi, D., Wang, B., Shen, F., Fu, S., Tang, C., Pan, S. "Effect of grain size on fatigue-crack growth in 2524 aluminum alloy." ***International Journal of Fatigue***, Vol. 24, (2016), 9-16, https://doi.org/10.1016/j.ijfatigue.2015.11.011.
23. Yang, B., Wu, M., Li, X., Zhang, J., Wang, H. "Effects of cold working and corrosion on fatigue properties and fracture behaviors of precipitate strengthened Cu-Ni-Si alloy." ***International Journal of Fatigue***, Vol. 116, (2018), 118-127, https://doi.org/10.1016/j.ijfatigue.2018.06.017.
24. Gavras, A.G., Lados, D.A., Champagne, V.K., Warren, R.J. "Effects of processing on microstructure and fatigue crack growth mechanisms in cold-spray 6061 aluminum alloy." ***International Journal of Fatigue***, Vol. 110, (2018), 49-62, https://doi.org/10.1016/j.ijfatigue.2018.01.006.
25. Shiozaki, T., Tamai, Y., Urabe, T. "Effect of residual stresses on fatigue strength of high strength steel sheets with punched holes." ***International Journal of Fatigue***, Vol. 80, (2015), 324-331, https://doi.org/10.1016/j.ijfatigue.2015.06.018.
26. Ritchie, R.O., Suresh, S. "Some considerations on fatigue crack closure at near-threshold stress intensities due to fracture surface morphology." ***Metallurgical Transaction A***, Vol. 13, (1982), 937-940, https://doi.org.ezproxy.ugm.ac.id/10.1007/BF02642409.
27. Suresh, S. "Fatigue crack deflection and fracture surface contact: micromechanical models." ***Metallurgy Transaction A***, Vol. 16, (1985), 249-260, https://doi.org.ezproxy.ugm.ac.id/10.1007/BF02815306.
28. Ilman, M.N., Sriwijaya, R.A., Muslih, M.R., Triwibowo, N.A., Sehono. "Strength and fatigue crack growth behaviours of metal inert gas AA5083-H116 welded joints under in-process vibrational treatment." ***Journal of Manufacturing Processes***, Vol. 59, (2020), 727–738, https://doi.org/10.1016/j.jmapro.2020.10.035.
29. Newman Jr, J.C. "The merging of fatigue and fracture mechanics concepts: a historical perspective." ***Progress in Aerospace Sciences***, Vol. 34, (1998), 347-390, https://doi.org/10.1016/S0376-0421(98)00006-2.
30. Matos, P.P., Moreira, P.P., Pina, J.C. "Residual stress effect on fatigue striation spacing in a cold-worked rivet hole." ***Theoretical and Applied Fracture Mechanics***, Vol. 42, (2004), 139-148, https://doi.org/10.1016/j.tafmec.2004.08.003.
31. Yan-li, W., You-li, Z., Shui, H., Han-xiao, S., Yong, Z. "Investigation on fatigue performance of cold expansion holes of 6061-&6 aluminum alloy." ***International Journal of Fatigue,*** Vol. 95, (2017), 216-228, https://doi.org/10.1016/j.ijfatigue.2016.10.030.
32. Muñoz-Cubillos, J., Coronado, J.J., Rodríguez, S.A. "Deep rolling effect on fatigue behavior of austenitic stainless steels." ***International Journal of Fatigue***, Vol. 95, (2017), 120-131, https://doi.org/10.1016/j.ijfatigue.2016.10.008.

Persian Abstract

چکیده

در مطالعه حاضر ، تأثیر نورد سرد و زمان بازپخت بر رفتار انتشار ترک خستگی آلیاژ آلومینیوم AA5052 بررسی شد .ورق آلیاژ تحت کاهش نوردهای مختلف ، به طور مثال ، 0 ، 15 ، 30 و 45 درصد نورد سرد شد .سپس45 ٪ نمونه نورد شده در دمای 370 درجه سانتیگراد تحت زمانهای مختلف پخت ، به مدت 2 ، 4 و 6 ساعت آنیل شد . تکاملهای ریزساختار پس از تیمارهای نورد سرد و بازپخت نیز با استفاده از میکروسکوپ نوری مورد بررسی قرار گرفت در حالی که رفتار انتشار ترک خستگی با استفاده از تست خستگی مشخص شد .نتایج نشان داد که دانه های کاملاً کشیده با افزایش کاهش نورد مشاهده شدند .ریزساختارهای کشیده به دلیل تبلور مجدد در حین عملیات بازپخت به ساختارهای هم محور تغییر یافت .عمر خستگی با افزایش کاهش نورد به شدت کاهش یافت اما با افزایش زمان پخت به طور قابل توجهی افزایش یافت .عمر خستگی آلیاژ با نورد سرد تا45 ٪93 ٪ کاهش می یابد .از طرف دیگر ، عمر خستگی45 ٪ نمونه های نورد شده در زمان حرارت دادن در دمای 370 درجه سانتیگراد به مدت 6 ساعت به میزان قابل توجهی 412 درصد افزایش یافت .

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 6 journal pages (including figures and tables). For the additional to 6 journal pages, there will be page charges. It is advised to keep papers limited to 3500 words.

Page Charges for Papers More Than 6 Pages (Including Abstract)	
For International Author ***	**$55 / per page**
For Local Author	**100,000 Toman / per page**

AUTHOR CHECKLIST

- Author(s), bio-data including affiliation(s) and mail and e-mail addresses).
- Manuscript including abstracts, key words, illustrations, tables, figures with figure captions and list of references.
- MS Word files of the paper.

International **J**ournal of **E**ngineering, P.O. Box 13145-1659, Tehran, I.R. Iran
Materials and Energy Research Center, 5 Ahuramazda St., Alvand Ave., Argentina Sq., Tehran, Islamic Republic of Iran
www.merc.ac.ir - www.emerc.ac.ir

THOMSON REUTERS

WEB OF SCIENCE

Scopus®

CAS®
A DIVISION OF THE
AMERICAN CHEMICAL SOCIETY

Google
Scholar