

International Advanced Researches and Engineering Journal

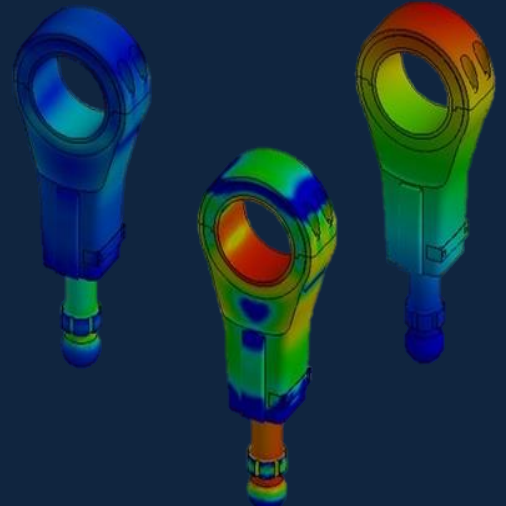
e-ISSN
2618-575X



$$F=ma$$

$$E=mc^2$$

$$\int \frac{dy}{dx} dt$$



Volume	Issue
05	01

April, 2021



e-ISSN: 2618-575X

Available online at www.dergipark.org.tr/en

INTERNATIONAL ADVANCED RESEARCHES
and
ENGINEERING JOURNAL

Journal homepage: www.dergipark.org.tr/en/pub/iarej

International
Open Access



Volume 05
Issue 01
April, 2021

International Advanced Researches and Engineering Journal (IAREJ) is a double-blind peer-reviewed and publicly available online journal that has Editorial Board (<https://dergipark.org.tr/en/pub/iarej/board>). The editor in chief of IAREJ welcomes the submissions that cover theoretical and/or applied researches on **Engineering** and related science with Engineering. The publication language of the Journal is **English**. **Writing Rules** are given in Author Guidelines (<https://dergipark.org.tr/en/pub/iarej/writing-rules>). IAREJ publishes **original papers** that are research papers and technical review papers.

IAREJ publication, which is **open access**, is **free of charge**. There is no article submission and processing charges (APCs).

IAREJ is indexed & abstracted in:

Crossref (Doi beginning: 10.35860/iarej.xxxxxx)
Directory of Open Access Scholarly Researches (ROAD)
Directory of Research Journals Indexing (DRJI)
EBSCO
Google Scholar
Index Copernicus (ICI Journal Master List)
J-Gate
TUBITAK ULAKBIM TR Dizin (TR index)
WorldCAT

Authors are responsible from the copyrights of the figures and the contents of the manuscripts, accuracy of the references, quotations and proposed ideas and the Publication Ethics (<https://dergipark.org.tr/en/pub/iarej/page/4240>).

All rights of the issue are reserved by International Advanced Researches and Engineering Journal (IAREJ). IAREJ also allows the author(s) to hold the copyright of own articles.

©

IAREJ

15 April 2021



This is an open access issue under the CC BY-NC license (<http://creativecommons.org/licenses/by-nc/4.0/>).



e-ISSN: 2618-575X

Available online at www.dergipark.org.tr/en

INTERNATIONAL ADVANCED RESEARCHES
and
ENGINEERING JOURNAL

Journal homepage: www.dergipark.org.tr/en/pub/iarej

International
Open Access



Volume 05
Issue 01
April, 2021

Table of Contents

Research Articles	Pages
Finite difference method for electric field optimization in high voltage power transformer bushings using engineering simulation and 3D design program <i>Nihat PAMUK</i>	01-07
Water cooled PV panel efficiency in Osmaniye environment <i>Halil EROL, Mahmut UÇMAN, Zehan KESİLMİŞ</i>	08-13
Silicon dioxide thin films prepared by spin coating for the application of solar cells <i>İmran KANMAZ, Abdullah ÜZÜM</i>	14-18
Selective harmonic elimination in multi-level inverters by using neural networks <i>Özkan AKIN, İbrahim ÖZER, Halil ÜNLÜ</i>	19-25
Effect of aperture averaging on four petal Gaussian beams in atmospheric turbulence <i>Mert BAYRAKTAR</i>	26-30
Analysis of whether news on the Internet is real or fake by using deep learning methods and the TF-IDF algorithm <i>Tilbe KORKMAZ, Ali ÇETİNKAYA, Hakan AYDIN, Mehmet Ali BARIŞKAN</i>	31-41
Recombinant production of Thermus aquaticus single-strand binding protein for usage as PCR enhancer <i>Özlem KAPLAN, Rizvan İMAMOĞLU, İskender ŞAHİNGÖZ, İsa GÖKÇE</i>	42-46
Dose response of gluconic acid doped Fricke gels irradiated with X-rays <i>Serkan AKTAŞ, Özlem KORKUT, Mustafa Erdem SAĞSÖZ</i>	47-52
GMDH-type neural network-based monthly electricity demand forecasting of Turkey <i>Ali Volkan AKKAYA</i>	53-60
Mechanical characterization of pack-boronized AISI 4140 and AISI H13 steels <i>Dilek ARSLAN, Selda AKGÜN</i>	61-71
Experimental investigation and numerical verification of Coanda effect on curved surfaces using co-flow thrust vectoring <i>Emre KARA, Hüdayi ERPULAT</i>	72-78
Modelling of cutting parameters for Nilo 36 superalloy with machine learning methods and developing an interactive <i>Gültekin BASMACI, İsmail KIRBAŞ, Mustafa AY</i>	79-86
Analysis and balancing of assembly line in a machine molding factory <i>Esra CAN, Adalet ÖNER</i>	87-96
An investigation on the flexural behaviour of RC beams wrapped with CFRP <i>İlknur DALYAN, Bilge DORAN</i>	97-105

Comparison of the effect of foundation analysis methods on structural analysis results of tall buildings <i>Ömer ÖZER, Bahadır YÜKSEL</i>	106-112
--	---------

Review Articles	Pages
Comparison of conventional high speed railway, maglev and hyperloop transportation systems <i>Mehmet Nedim YAVUZ, Zübeyde ÖZTÜRK</i>	113-122
An analysis of content-based image retrieval <i>Hakan KOYUNCU, Manish DIXIT, Baki KOYUNCU</i>	123-141



Research Article

Finite difference method for electric field optimization in high voltage power transformer bushings using engineering simulation and 3D design program

Nihat Pamuk ^{a,*} 

^aZonguldak Bülent Ecevit University, Department of Electric Electronic Engineering, Zonguldak, 67100, Turkey

ARTICLE INFO

Article history:

Received 07 July 2020

Revised 29 September 2020

Accepted 08 October 2020

Keywords:

Electric field distribution

Finite difference method

Optimization

Power transformer bushing

3D design

ABSTRACT

The electric field optimization minimizing the field strength on an electrode surface and providing its uniformity is important in designing high voltage power transformer bushings and other apparatus from the viewpoint of efficient utilization of the electric field space. The high voltage power transformer bushing with cylinder electrode system has been designed and tested in this investigation. It was found that the insulation method of the cylinder electrode was the most important factor to lower streamer initiation voltage. The optimized design uses both internal and external elements for electric stress grading at critical parts of the bushing. Applying optimization theory based on charge simulation method, the author developed a computation program for electric field automatic optimization in 3D dielectric axisymmetric field. The results of the computation realized some excellent electrode profiles with uniform electric field distribution. Moreover, the discrepancy from the electric field uniformity on 3D dimensional profile caused by applying it to an axisymmetric electrode was discussed. Then a new electrode with uniform field distribution was obtained by using the computation program for optimization.

© 2021, Advanced Researches and Engineering Journal (IAREJ) and the Author(s).

1. Introduction

A number of fine results have been obtained by some numerical electric field computations methods that have been greatly improved during recent years. Numerical electric field computation is now indispensable to the designing of high voltage power transformer bushings and other insulated apparatus from the viewpoint of reliability and reduction in size of apparatus [1]. In the case of designing SF₆ gas insulated apparatus, the field calculation has special meaning of importance because the insulation characteristics of SF₆ gas greatly depend on the electric field strength [2]. As well as to analyze the field a number of efforts have recently been made to obtain optimum electrode profiles having minimum and uniform electric field [3, 4].

The main purpose of electric field optimization is to obtain uniform electric field distribution as well as to minimize the field strength on the electrode surface [5]. This enables one to make the best use of the electric field space, reducing the size of high voltage power transformer

bushings and other insulated apparatus. Electric field calculations by numerical techniques is to use different numerical techniques to find electric field distributions, which are inevitable tool in various electricity concerned technologies [6]. Up to now, several 3D dimensional electrode profiles with quasi-uniform electric field have been obtained using conformal mapping and experiments. They are noted as a finite difference method profile and finite element method profile [7, 8].

The examined electrode systems are modeled as 2D using axial symmetry. In this approach, it is considered that 3D electric field distribution is obtained by rotating 2π radians around the symmetry axis of the models. Among them is the finite difference method profile which has a completely uniform electric field along the electrode surface. However, this complete electric field uniformity will disappear if this profile is applied to a 3D dimensional axisymmetric electrode.

Huang et al. and Hyouk et al. calculated electric field distribution of the two dimensional finite difference

* Corresponding author. Tel.: +90-372-291-2617

E-mail addresses: nihatpamuk@beun.edu.tr; nihatpamuk@gmail.com (N. Pamuk)

ORCID: 0000-0001-8980-6913 (N. Pamuk)

DOI: 10.35860/iarej.765360

This article is licensed under the CC BY-NC 4.0 International License (<https://creativecommons.org/licenses/by-nc/4.0/>).

method profile and 3D dimensional axisymmetric finite difference method profiles to evaluate the discrepancy from the uniformity for the latter profiles. Furthermore, by using the computation program for high voltage electric field automatic optimization, a new uniform field electrode profile was obtained as a replacement of the 3D dimensional finite difference method profile. Other optimum 3D dimensional axisymmetric electrode profiles are also obtained which are enclosed with metallic coaxial cylinder [9, 10].

Most of the studies found that there was electrostatic field calculation and optimization structure in power transformer bushing insulation, while some studies such as [11, 12] found that there was condition assessment of high voltage bushing with solid insulation structure. Some studies such as [13, 14] have found that electromagnetic forces and losses computation and design of power transformer bushings. By applying theory of electric field optimization based on charge simulation method, the author developed a computation program for electric field automatic optimization of an electrode profile with 3D dielectrics in axisymmetric field.

The purpose of this paper is to make the radial stiffness characteristic of the given 3D rubber bushing model meet the target stiffness curve by using the optimization method presented in this paper. In order to determine the stiffness curve of the bushing, nonlinear finite model was defined using boundary conditions.

2. Basic Equations for the Method of Optimization

The charge simulation method is made up of the following basic equation. To simplify of discussion, the author consider one dielectric field problem.

$$[P] \cdot \{Q\} = \{\phi\} \quad (1)$$

In equation 1, P is a potential coefficient matrix determined by coordinates of contour point and charge point. Q is a charge vector. ϕ is a potential vector of contour points. In general, P and ϕ are given, thus providing the value of Q. From the value of Q, the author can calculate the electric field strength. [15, 16] proposed that electric field should be improved by properly arranging several new charges (optimizing charges) in the field obtained by charge simulation method. Let Q' be an optimizing charge vector.

$$[P] \cdot \{Q\} + [P'] \cdot \{Q'\} = \{\phi\} \quad (2)$$

In equation 2, P' is a matrix of varied part of potential coefficients determined by arrangement of new optimizing charges. Combining with the given value Q', the value Q obtained by equation 2 determines the electric field distribution along a new electrode. The profile of a new electrode is obtained as a locus of equipotential line of electrode potential. If the electric field on a new electrode surface does not distribute as uniformly as expected, the

above computation is iterated by giving new Q' to equation 2 [17, 18, 19].

This computation is iterated until uniform field distribution of electrode surface is obtained. For a good convergence and satisfactorily uniform potential distribution, it is important to determine the value of Q' appropriately. This is discussed in the following section. On the other hand, the basic equation of the charge simulation method in two electrostatic dielectric field may be generally expressed by equation 3.

$$\begin{bmatrix} P \\ F \end{bmatrix} \cdot [Q] = \begin{bmatrix} \phi \\ 0 \end{bmatrix} \quad (3)$$

In equation 3, F is an electric field coefficient matrix determined by dielectric boundary points. Like equation 2, equation 3 is modified by arranging new optimizing charges Q', as follows equation 4.

$$\begin{bmatrix} P \\ F \end{bmatrix} \cdot [Q] + \begin{bmatrix} P' \\ F' \end{bmatrix} \cdot [Q'] = \begin{bmatrix} \phi \\ 0 \end{bmatrix} \quad (4)$$

In equation 4, F' is a matrix of varied part of electric field coefficients determined by arrangement of optimizing charges. Like in one dielectric field, electrode surface electric field can be optimized also in two dielectric field, using equation 4.

3. Automatic Optimization Technique

The author succeeded in obtaining the optimized electric field distribution by using iterative computation by arranging vector (Q') automatically. In each iteration step, Q' is given to correct local maximum and minimum of the electric field distribution along the electrode profile [20]. This method is described along with a model for consideration in Figure 1.

In Figure 1, A-G is an electrode profile to which optimization is to be applied. Other parts of the profile do not vary through the iterative computation since these points are always fixed as the contour points. Optimizing charge vector (Q'₁) given at an initial time is arranged at a positions shown by symbol ●. The electrode profile calculated (solid curve) is identical to the equipotential surface. The electric field distribution in A-G is expressed by solid curve shown in Figure 2.

This curve provides local maximum points at B, D, F and local minimum points at C and E. To correct this electric field distribution, new charges (optimizing charges) are arranged at the vicinity of A, C and E. (shown by symbol ◻) Figure 3 shows a detail of arrangement of an optimizing charge. Taking point C as an example, the author arrange the optimizing charge Qc' at point S away from point C by distance ℓ_{cs} . ℓ_{cs} may be expressed by equation 5.

$$\ell_{CS} = k \cdot \text{Min}(\ell_{BC}, \ell_{CD}) \quad (5)$$

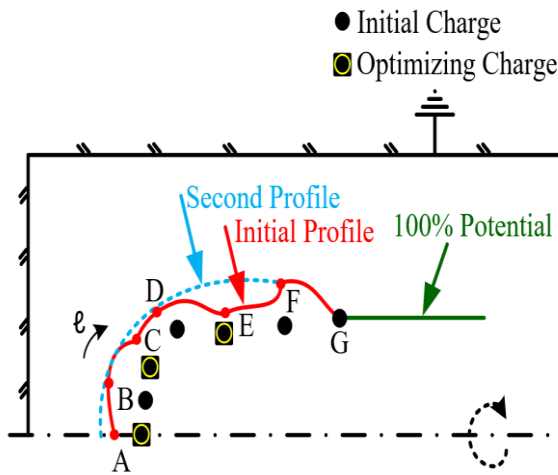


Figure 1. Optimization of electrode profile

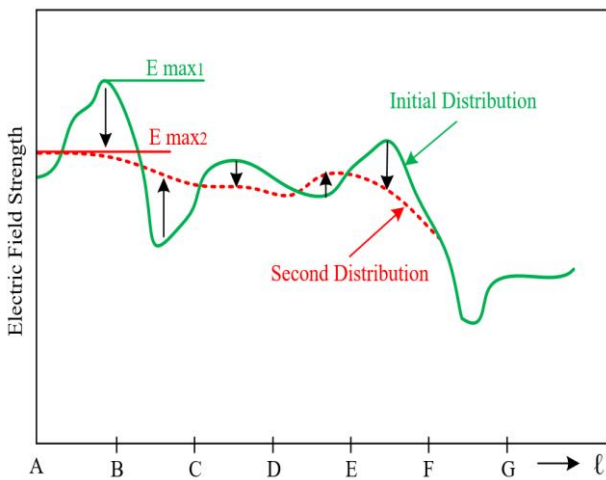


Figure 2. Electric field strength distribution

In equation 5, k gives an appropriate value of $1 \leq k \leq 2$. The value of Q_c' is determined by equation 6 using potential value jV_{CB} and jV_{CD} , which are given by a function f in equation 7, where E_j is an electric field strength on next following point j .

$$Qc' = (*) \cdot \frac{v}{j} (P_{cj}^{-1} \cdot jV_{cj}) \quad (j = B, D) \quad (6)$$

$$jV_{cj} = f \cdot (|E_j - E_c|) \quad (j = B \text{ or } D) \quad (7)$$

In equation 6, P_{cj} is a potential coefficient, and $*$ is the relaxation factor which is introduced in order to get smooth convergence of iterative computation. The value of $*$ depends on the electrode profile. Let (Q'_2) be a newly given vector by the above procedure, the equipotential surface (new electrode surface) shown by dotted line in Figure 1 is obtained. The corresponding electric field distribution is shown by dotted line in Figure 2. The above calculation process is automatically iterated and the electrode profile with the optimized electric field distribution is obtained by judging the convergence from a reduction rate of the maximum electric field and an achievement of uniformity of electric field distribution.

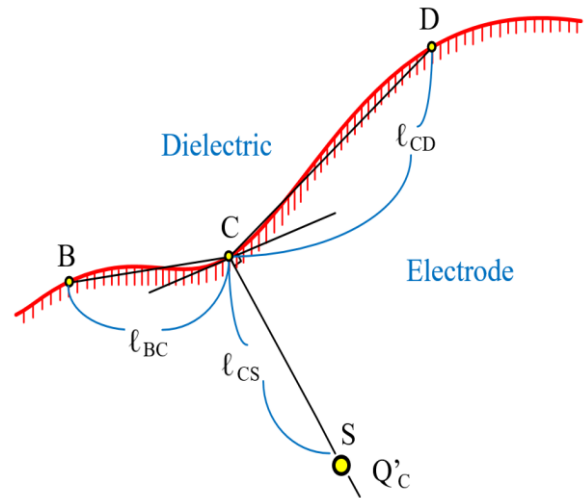


Figure 3. Arrangement of optimizing charge

4. Examples of Optimized Electrode Profile

This section gives some calculation examples using above mentioned computation program for optimization. Figure 4 shows an example of the end profile of high voltage conductor in a ground potential cylindrical enclosure. Figure 4 indicates an electrode profile at the initial stage with optimizing charge positions, while Figure 5 shows those as a result of automatically iterative calculation of five times.

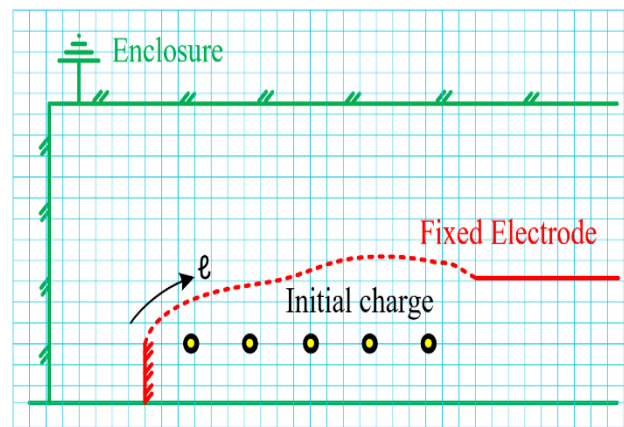


Figure 4. Initial electrode profile

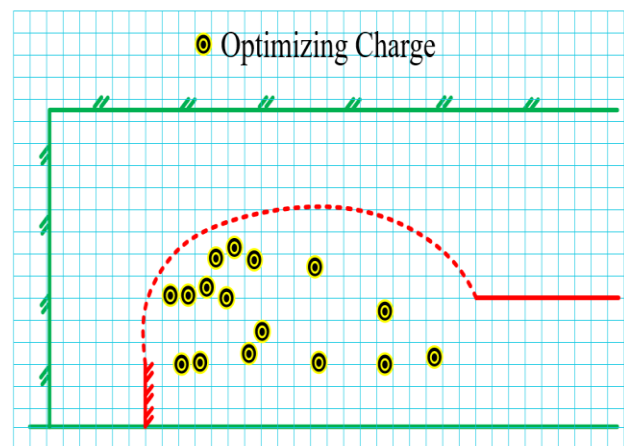


Figure 5. Optimized electrode profile

Figure 6 shows a variation of electric field distribution. The iterative calculation reveals an obvious reduction of the maximum electric field strength and a uniformity of electric field distribution.

Figure 7 shows further example of optimized profile of another type conductor end in grounded cylindrical enclosure, and the electric field distribution along it. Optimized profile and electric field distribution obtained with and without the insulating cylinder in high voltage power transformer bushing are given in Figure 8 for comparative study.

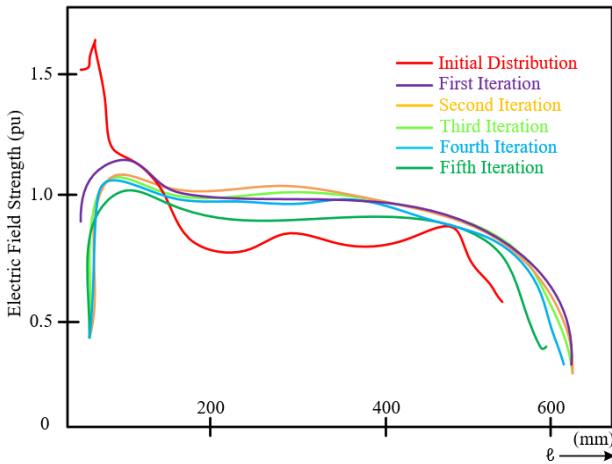


Figure 6. Electric field distribution

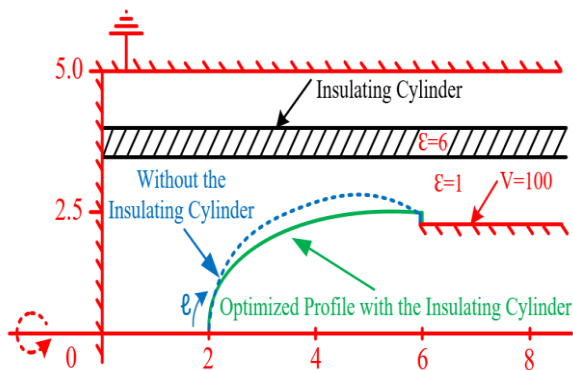


Figure 7. Optimized profile with and without insulating cylinder

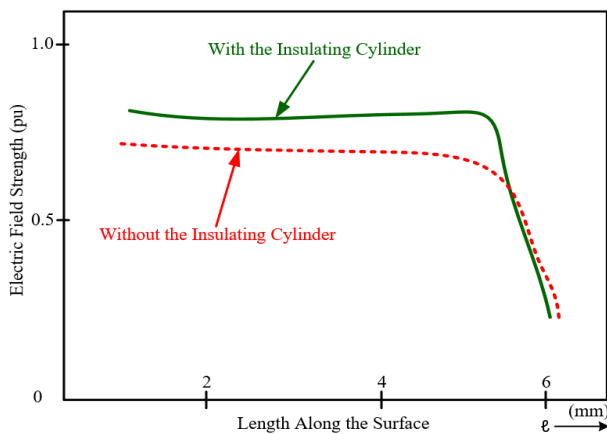


Figure 8. Electric field distribution with and without insulating cylinder in high voltage power transformer bushing

5. 3D Dimensional Electrode Profile

3D dimensional electrode profile ($\pi/2$ dimensional profile) against infinite plane may be expressed as follows by using parameter φ . (Point A is assumed as an origin.)

$$x = -2 \left[\sin\varphi - \ell \tan\left(\frac{\varphi}{2} + \frac{\pi}{4}\right) \right] \quad (\varphi = 0 - \frac{\pi}{2}) \quad (8)$$

$$y = 2[1 - \cos\varphi] \quad (\varphi = 0 + \frac{\pi}{2}) \quad (9)$$

$$z = 2 \left[\sin\varphi + \ell \tan\left(\frac{\varphi}{2} - \frac{\pi}{4}\right) \right] \quad (\varphi = 0) \quad (10)$$

Figure 9 illustrates 3D dimensional electrode profile with 100% potential. The grounded infinite plane electrode is located at $x = 2-\pi$, $y = 2+\pi$ and the gap length is π as $z \rightarrow \infty$. Arrows in Figure 9 drawn by computer show electric field vectors on the 3D dimensional electrode profile.

The electric field strength is calculated by using charge simulation method. The results reveal that the electric field distribution on 3D dimensional electrode profile is completely uniform. The author apply the 3D dimensional electrode profile to an axisymmetric electric field distribution. One may consider an end portion profile of a rod electrode. The problem is, however, that this 3D dimensional electrode profile loses the uniformity of the electric field distribution along the profile by influence of axial curvature and that electric field strength increases as going apart from the axis. By using charge simulation method, the author calculated the discrepancy of the electric field uniformity. Figure 10 shows the result as a function of 3D coordinate, where (x_0, y_0, z_0) is the coordinate of edge point A assuming the axis center of the infinite plane as origin.

Minimum electric field strength E_{min} is obtained on the center axis of the 3D dimensional electrode profile, while the maximum electric field strength E_{max} is obtained on point A. Figure 10 reveals that the discrepancy from the uniformity of electric field is 10% or more if $|x_0/y_0/z_0| < 5$ and 5% or more if $|x_0/y_0/z_0| < 10$. In other words, 2D dimensional electrode profile directly applied to 3D dimensional axisymmetric field cannot be used as an insulation breakdown test. The fatal problem is that insulation breakdown is expected to occur at edge point A of this electrode.

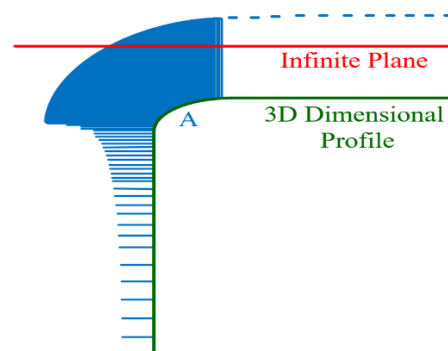


Figure 9. 3D dimensional profile and electric field distribution

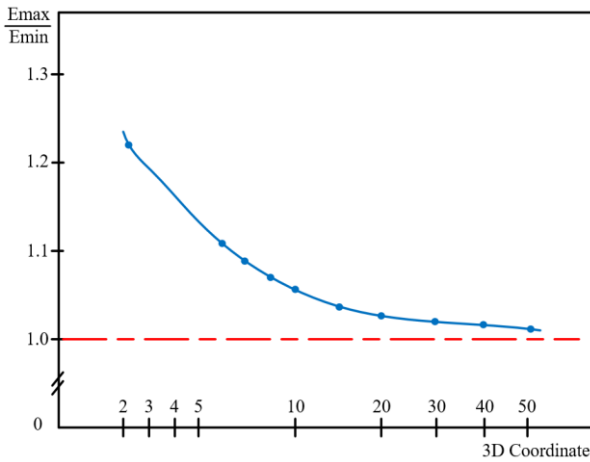


Figure 10. Discrepancy from the electric field uniformity

6. Optimized Profile in Axisymmetric Field

First of all, the parameter values of the elastoplastic transformer bushing model were obtained by using the finite difference analysis method. With the experiments applied in the design phase, the bandwidths of the hysteresis curve were determined by fitting the curve model obtained as a result of the least squares regression method. The software values of the hysteresis curve obtained as a result of the finite difference analysis take into account the damping force and curve bandwidth values. Thus, both viscoelastic and elastoplastic components are represented. In addition, the slope value of the hysteresis curve shows the decreased stiffness value due to the amplitude effect and its behavior is represented by the elastoplastic component value.

The rubber structured transformer bushing sample was modeled based on the geometric structure of the finite difference analysis. It is ensured that the boundary condition value of the finite difference analysis model, the constant condition value on the outer flange of the transformer bushing and the displacement value applied to the inner rod element of the bushing at the central point are the same as the experimental conditions. Rotational hardness information was obtained by applying torque from 0 Nmm to 6000 Nmm in the x direction and from 0 Nmm to 600 Nmm in the y and z direction, respectively. Rotational stiffness tests were not conducted due to the limited number of experimental devices. For this reason, the values obtained by simulation based on finite difference analysis values were used.

The electric field on the 2D dimensional electrode profile applied to 3D dimensional axial symmetry does not meet requirements of uniform electric field distribution. The author calculated a new uniform electric field electrode profile by using charge simulation program for electric field automatic optimization. Figure 11 shows a part of the results and has been obtained by iterative computation with a starting value of 3D dimensional electrode profile when $|x_0/y_0/z_0| = 4$. The distortion of surface electric field strength is less than 0.4% if $|x| < 3.5$ and the strength slightly reduces

in the vicinity of the point A.

This new electrode profile is most appropriate to one for an insulation breakdown testing because uniform electric field strength is realized on a wide electrode area and the electric field decreases as going apart from the uniform field region. The author discussed the case where a high voltage rod electrode is in a cylindrical grounded enclosure. The author obtained uniform electric field distribution of the end portion from many trials with various distances from the enclosure wall. Figure 12 shows the result, and D_i is a diameter of the rod electrode D_0 an inner diameter of the grounded enclosure.

Also, G is a gap length at the rod end portion center. Assuming that $G = 1$ and $D_i = 20$, the author obtained the optimum electrode profile as a function of D_0 . If $D_0 = 22.104$, the uniform electric field spreads over the entire electrode surface. If D_0 is larger, the uniform electric field is obtained only at the top end portion of the electrode. If $D_0 = \infty$, the profile is identical to the profile optimized from 3D dimensional electrode profile previously discussed. Figure 12 shows that there is almost no variation in the optimum profile if $D_0 \geq 30$, meaning that the grounded enclosure wall does not so much affect the electric field on the electrode surface.

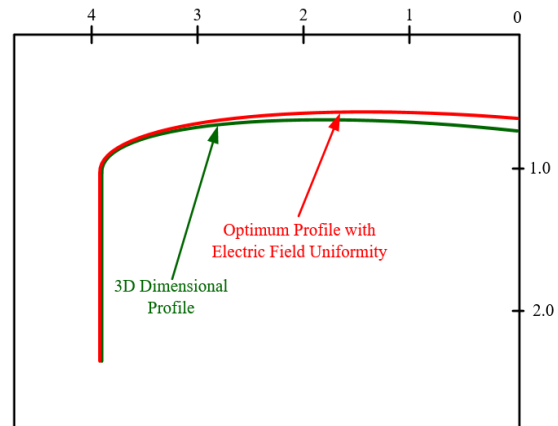


Figure 11. Axisymmetric 3D dimensional electrode profile with electric field uniformity

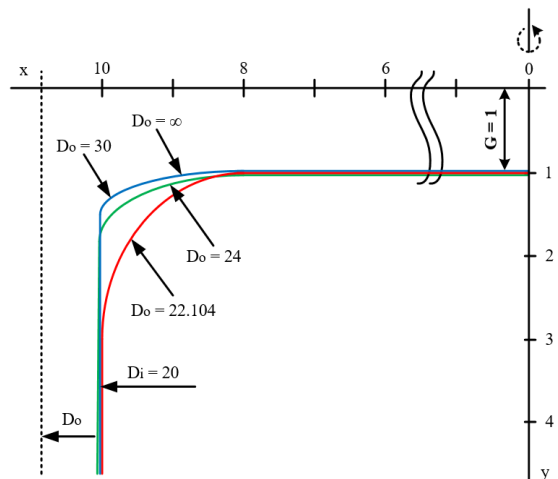


Figure 12. Optimized electrode profile with cylindrical enclosure

7. Conclusions

Satisfactory results were obtained from the computation program for electric field automatic optimization in high voltage power transformer bushing developed by applying charge simulation theory of electric optimization. In above mentioned simulation program, the main feature is that new charges are arranged consecutively in order to correct the variation in electric field distribution of the electrode surface, thus making it possible to achieve the electric field uniformity by automatic iterative computation. Electric field calculation using charge simulation method performed for a 2D dimensional electrode profile revealed that the uniformity of electric field strength distribution along it is completely obtained.

Extended 3D dimensional electrode profile to a 3D dimensional axisymmetric field does not have an electric field uniformity any longer. Deviation from the uniformity was calculated by charge simulation method, and it has been clarified that the smaller the electrode diameter and the larger the distortion electrode diameter. In axisymmetric 3D dimensional field, the computation program for electric field automatic optimization developed by the author realized a new electrode profile that provides a uniform electric field distribution. Such electrode is suitable for an insulation breakdown testing as a replacement of the 3D dimensional electrode profile. The computation for electric field optimization on a rod electrode in a grounded cylindrical enclosure provided the optimized electrode profile as a function of enclosure diameter. If the inner diameter of grounded enclosure is 30 or more there is almost no variation in optimized configuration for a rod electrode of a diameter of 20 for a gap length of 1.

Declaration

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article. The author also declared that this article is original, was prepared in accordance with international publication and research ethics, and ethical committee permission or any special permission is not required.

Author Contributions

N. Pamuk is responsible for all section of the study.

References

1. El-Makkawy, S.M. El-Dessouky, S.S., *Analytical Aspects in the Presence of Polymeric Materials in Field Gaps of HV Insulation System*, Proceedings of 8th International Symposium on Electrets (ISE 8), 06 August 2002, Paris, France, p. 905-910.
2. Weifang, J. Huiming, W. Kuffell, E., *Application of the Modified Surface Charge Simulation Method for Solving Axial-symmetric Electrostatic Problems with Floating Electrodes*, Proceedings of 4th International Conference on Properties and Applications of Dielectric Material, 1994, Brisbane, Qld, Australia, p. 28-30.
3. Abidaoun, H.S. Maather, A.I. Mohanad, H.A. Saad, Q.F., *Estimation and Plot of Electrical Field Using Finite Difference Method*, Second Engineering Scientific Conference College of Engineering, 16-17 December 2015, University of Diyala, p. 501-510.
4. Vahidi, B., Mohammadzadeh Fakh Davood A., *Application of charge simulation method to electric field calculation in the power cables*. Iranian Journal of Science & Technology, Transaction B, Engineering, 2006. **30**(B6): p 789-794.
5. Hu, R. Zhang, Z. Wang, S. Lu, Y. Liu, L. Zhu, S. Peng, Z., *Electric Field Optimization of Cast Resin Dry-Type Transformer Under Lightning Impulse*, IEEE Conference on Electrical Insulation and Dielectric Phenomena (CEIDP), 20-23 October 2019, Richland, USA, p. 556-559.
6. Biswanath, M., *Electric field calculations by numerical techniques*, Bachelor Thesis in Electrical Engineering 2009, National Institute of Technology Rourkela, India. p. 6-7.
7. Zhou, K., Ivanco, A., Filipi, Z., Hofmann, H., *Finite element based computationally efficient scalable electric machine model suitable for electrified powertrain simulation and optimization*. IEEE Transactions on Industry Applications, 2015. **51**(6): p 4435-4445.
8. Lee, K.H., Hong, S.G., Baek, M.K., Choi, H.S., Kim, Y.S., Park, I.H., *Alleviation of electric field intensity in high voltage system by topology and shape optimization of dielectric material using continuum design sensitivity and level set method*. IEEE Transactions on Magnetics, 2015. **51**(3): p. 1-4.
9. Huang, Y., Xu, Q., Tan, Q., Xie, N., *Optimization of electric field distributions in OVS with hybrid algorithm*. IEEE Sensors Journal, 2019. **19**(21): p 9748-9754.
10. Hyouk Lee, K., Geon Hong, S., Ki Baek, M., Soon Choi, H., Sun Kim, Y., Han Park, I., *Adaptive level set method for accurate boundary shape in optimization of electromagnetic systems*. COMPEL The International Journal for Computation and Mathematics in Electrical and Electronic Engineering, 2014. **33**(3): p. 809-820.
11. Du, B., Sun, H., Jiang, J., Kong, X., Yang, W., *Temperature dependent electric field distribution in ± 800 kV valve-side bushing insulation for a converter transformer*. High Voltage, 2020.
12. Subocz, J., Mrozik, A., Bohatyrewicz, P., Zenker, M., *Condition assessment of HV bushings with solid insulation based on the SVM and the FDS methods*. Energies, 2020. **13**(4): 853, p. 1-13.
13. Li, Z., Yujiao, Z., Guanteng, X., Jiansheng, Y., Lan, J., *Electrostatic field calculation and structure optimization of new shield ring of 1000 kV Nan-Jing loop transmission line*. Transactions on Electrical and Electronic Materials, 2020. p. 1-8.
14. Al-Abadi, A. Gamil, A. Schatzl, F., *Optimization of Magnetic Shunts Towards Efficient and Economical Power Transformers Design*, Proceedings of the 21st International Symposium on High Voltage Engineering, 28 November 2019, ISH2019, p. 15-26.
15. Cao, Y., Liu, X., Wang, E., Jin, L., Wang, G., *Electric field optimization design of a vacuum interrupter based on the tabu search algorithm*. IEEE Transactions on Dielectrics and Electrical Insulation, 2002. **9**(2): p 169-172.
16. Zhao, Y.N., Zhang, G.Q., Guo, Z.Z., Cheng, S., *The mathematical model of electrical field distribution in optical voltage transformer*. Procedia Engineering, 2012. **29**(1):

p.2661-2666.

17. Tan, Q., Xu, Q., Chen, L., Huang, Y., *A new method to improve internal electric field distributions of poekels OVS*. IEEE Sensors Journal, 2017. **17**(13): p 4115-4121.
18. Wang, S., Kang, J., *Shape optimization of BLDC motor using 3D finite element method*. IEEE Transactions on Magnetics, 2000. **36**(4): p. 1119-1123.
19. Okamoto, Y., Masuda, H., Kanda, Y., Hoshino, R., Wakao, S., *Convergence acceleration of topology optimization based on constrained level set function using method of moving asymptotes in 3D nonlinear magnetic field system*. IEEE Transactions on Magnetics, 2017. **53**(6): p. 1-4.
20. Ho, S.L., Chen, N., Fu, W.N., *A moving mesh embedded algorithm in finite element method for optimal design of electromagnetic devices*. IEEE Transactions on Magnetics, 2011. **47**(10): p. 2947-2950.



Research Article

Water cooled PV panel efficiency in Osmaniye environment

Halil Erol ^{a,*} , Mahmut Uçman ^b  and Zehan Kesilmiş ^c 

^aOsmaniye Korkut Ata University, Electrical and Electronics Engineering, Osmaniye 80000, Turkey

^bOsmaniye Technical and Vocational School, Osmaniye 80000, Turkey

^cAdana Alparslan Türkeş Science and Technology University, Electrical and Electronics Engineering, Adana 01250, Turkey

ARTICLE INFO

Article history:

Received 28 August 2020

Revised 29 December 2020

Accepted 14 January 2021

Keywords:

Electrical efficiency

Maximum power

Photovoltaic/thermal (PV/T)

PV panel

PV water cooling

Solar system

Thermal efficiency

ABSTRACT

In this study, the electrical and thermal efficiency of the natural convection water-cooled photovoltaic panel (PV) is compared with the standard PV module. PV modules are made up of polycrystalline solar cells. As it is known that, an increase in PV panel temperature results in a decrease in electrical efficiency. The aim of this study is to increase PV panel electrical efficiency. PV panel characteristic values such as air/PV panel temperature, solar radiation, voltage, current, and power are recorded for both panels to the computer simultaneously. The thermal and electrical energy performance of PV panels are analyzed comparatively. The results are presented in detail. The water circulation structure is mounted under the PV module by using a glue that has good heat conductivity. The structure contains an S-shaped pipe and a plate made of copper. The plate is used for better heat absorption from the PV panel which is mounted downside of the panel with glue. The efficiency of the PV module with having a proposed cooling system and normal PV module is analyzed. For the overall efficiency, it is observed that the water-cooled PV system is better than the standard PV module by % 6.2.

© 2021, Advanced Researches and Engineering Journal (IAREJ) and the Author(s).

1. Introduction

Demand for alternative energy resources is increasing nowadays. Due to environmental concerns, renewable energy resources are taking more attention in recent years. Solar and wind energy production is the most popular and used renewable energy resources [1]. Two method is used for energy harvesting from solar system. The one is direct method by using PV panels. Whereas, in the other method solar energy is converted into thermal energy. In Concentrated Solar Power (CSP) technology, deploying parabolic mirrors connected to molten salt energy storage is used. After then the electricity production is obtained by use of this thermal energy. These systems are generally large scale systems [2]. CSP solar technology is widely used in large power facilities due to its storage ability. The storage ability is an important aspect since demand and production of power time do not coincide every time. That is power produced cannot be consumed simultaneously.

Solar PV systems has high initial investments costs. These can be reduced by improving PV cell efficiency.

Efficiency increasing studies for PV panels can be divided into two areas. The one in PV cell semiconductor design technology. Sunlight harvesting of perovskites type solar cells are better than silicon solar cells. Perovskites type solar cells are attained power conversion efficiency of 27% [3]. Many papers have been written in intrinsic semiconductor PV design technology. For example, electricity production by using cadmium telluride (CdTe) photovoltaic module technology attained the lowest-cost electricity in the solar industry. This is obtained by alloying selenium into the CdTe absorber. As a result, PV cell efficiency is increased from 19.5% to 22.1% [4]. The second area is related with improving the working condition of PV cells. According to the PV cells inherent characteristics, PV cell efficiency can be increased, (1) by increasing the incident solar radiation falling on a PV cell, (2) by holding the temperature of PV cells at optimum temperatures, (3) by selecting an appropriate working point on I-V curve for maximum power harvesting. In this article 2nd method is conducted with a novel cooling

* Corresponding author. Tel.: +90 328 8251818; Fax: +90 328 8251818.

E-mail addresses: halilerol@osmaniye.edu.tr (H.Erol), uc-ay@hotmail.com (M. Uçman), zkesilmis@atu.edu.tr (Z.Kesilmiş)

ORCID: 0000-0001-6171-0362(H.Erol), 0000-0002-5348-450X (M. Uçman), 0000-0002-5781-9450(Z.Kesilmiş)

DOI: 10.35860/iaiej.787168

This article is licensed under the CC BY-NC 4.0 International License (<https://creativecommons.org/licenses/by-nc/4.0/>).

approach. Different PV cooling techniques are used in literature [5, 6]. PV modules can not convert whole solar radiation into electricity. Only around 20-25 % of solar radiation could be converted. The rest of the energy is converted into heat [7, 8]. The temperature of the PV cell increases gradually by environmental conditions. This results in a decrease in PV cell electrical efficiency [9, 10]. For this reason, "PV cooling" is one of the main methods to increase electrical power generation in PV modules. [9].

Recycled cooling water can be used as hot water demands. For example, it might be used as hot water supply for homes [10].

Lifetime and Efficiency of PV cell is important. In literature, the electrical efficiency of PV cell is increased between 15% to 23% [11]. The effect of temperature raise is higher in polycrystalline and monocrystalline silicon solar cells which is about 0.45%/°C, with respect to amorphous silicon cells which is about 0.25%/°C. Moreover, amorphous silicon has lower thermal coefficient with respect to the other technologies. For this reason, it is widely used in thermal cogeneration applications [12, 13]. To reduce cell temperature different cooling techniques are used. These can be natural or forced type cooling. Also, different coolants are used such as air, water, oil and some chemical compounds [9, 14].

Different cooling techniques are used in PV cell cooling. Some of them are air based cooling (air channel, air gap), liquid based cooling (water flow, jet impingement, liquid immersion, submerging), phase change materials based (conductive) cooling. Matias et al., studied water cooling of PV panel with different coolant flow rate and compared the results [15]. They obtained efficiencies ranging from 5.30 % to 22.69% depending on the flow rate.

In this study PV cell cooling with naturel flow with a novel approach is experimentally analyzed. Cooling mechanism is mounted on back of PV panel. It consists of mainly two parts, copper sheet and copper pipe which is mounted on the sheet in S model. The gap between the pipes are approximately 2 cm. The gap between pipes is minimized to obtain a better cooling effect. Water circulation is achieved by natural flow. PV panel cooling performance and the overall heat transfer can be increased by increasing the circulation capacity. However, increasing water consumption used in cooling will increase cost of power production. This results in reduction of efficiency of overall system.

The experimental set up contains a PV module with the designed cooling system, two water tanks, and a DC load as shown in Figure 1. One of the tanks is used for hot water the other is for cold water storage. Water circulates from the cold water tank to the hot water tank by natural convection. The purpose of the DC load is to sample the I/V curve of the PV module. PV module is loaded with DC

load by varying the load resistance to obtain I/V curve of the module.

In experiments, PV cell temperature, solar radiation, environment temperature, cooling water temperature, voltage and current values of panel are recorded. Two Tommatech brand 10W polycrystalline PV modules are used in experimental analysis. One of the panels is equipped with a copper plate that has pipes. The copper pipes are attached on a sheet of copper with a thickness of 0.5 mm which adhered to the back of the PV module. The second panel has no extra cooling equipment.

2. Solar Insolation Conditions in Osmaniye Environment for PV Energy Production

In this section insolation duration and sun radiation for Osmaniye province is tabulated. The data presented here is obtained from General Directorate of Meteorology of Turkish Republic website.

The monthly average solar radiation in the Osmaniye environment is above the Turkey average. Monthly global radiation for a long time period, 2004-2018, in Osmaniye environment is shown in figure 2. Also, average insolation duration in Osmaniye is above the Turkey average. Monthly insolation duration is given in figure 3. The whole data is obtained from General Directorate of Meteorology website. Longer insolation time and higher radiation means better PV output power. In Osmaniye province, the sun radiation and insolation duration suggest that PV energy production in this environment is efficiently sustainable.

3. Mathematical Model of Single PV Module

Although many models have been proposed in the literature, two types of PV cell models stand out. The first of these is the single diode model as shown in figure 4. The second is two diode models as shown in figure 5.

In this article, a single diode model is preferred due to its ease of use. The I/V characteristic of the model has been obtained by assuming constant temperature and solar insolation conditions. I/V relationship of the single diode PV can be defined as follows:



Figure 1 Experimental setup of the developed PV system.

$$I_D = I_0 \left(e^{\frac{(V_D - R_s I_{pv})}{nV_T}} - 1 \right) \text{ where } V_T = kT/q \quad (1)$$

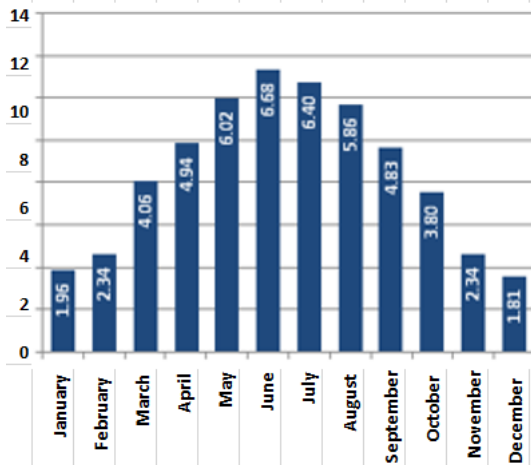


Figure 2. Global radiation for Osmaniye center. (kWh/m².day)

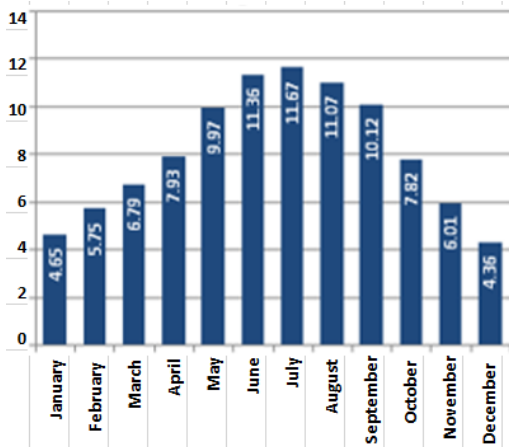


Figure 3. Insolation duration for Osmaniye city, hourly basis

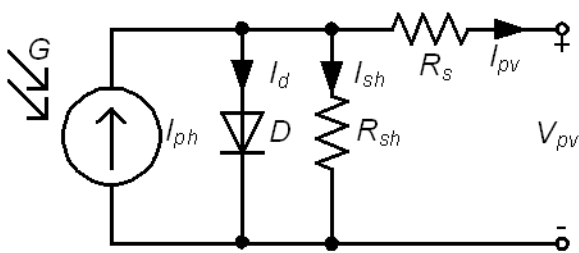


Figure 4. One diode equivalent model of PV panel.

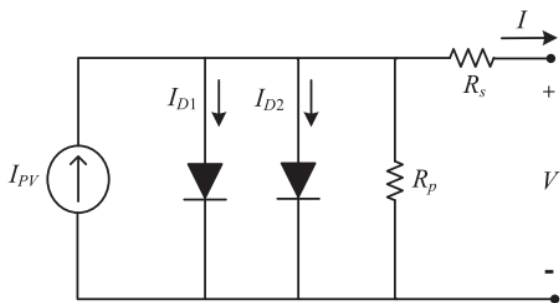


Figure 5. Two diode equivalent model of PV panel.

R_{sh} in the model stands for the p-n junction leakage current in the solar cell. There is a metal base between the semiconductor layers. This metal base has a contact resistance. This resistance is modeled in design with the existence of R_s [16]. The diode D assumed to have I-V characteristic of the Shockley diode model as shown in equation (1), where I_0 is reverse bias saturation current of the diode, I_D is the current that flow through diode D , V_D is the voltage across the diode D , k is the Boltzmann constant ($1.3806503 \times 10^{-23}$ J/K), n is the diode ideality factor, the thermal voltage is represented by V_T , operating temperature of cell in degree Kelvin is represented by T and q is the electron charge ($1.60217646 \times 10^{-19}$ C).

PV module load current is:

$$I_{pv} = I_{ph} - I_{sh} \left(e^{\frac{q(V_{pv} - R_s I_{pv})}{NKT}} - 1 \right) - \frac{(V_{pv} - R_s I_{pv})}{R_{sh}} \quad (2)$$

Where I_{ph} is the main current produced by solar cell, the solar cell terminal current is symbolized by I_{pv} , the solar cell terminal voltage is symbolized by V_{pv} , the equivalent parallel resistance is represented by R_{sh} and the equivalent series resistance is represented by R_s . The resistances R_{sh} and R_s in equation (2) can be calculated by iteration. In iteration the constrained that the maximum power obtained from model should conform with peak power from the datasheet at MPP [17]. Table 1 shows some assumptions made for parameter estimation of PV panel.

At Maximum Power Point, voltage (V_{mp}) and current (I_{mp}) is used for calculation of the maximum output power (P_{MP}), as follows:

$$P_{mp} = I_{mp} V_{mp} \quad (3)$$

The technical specification and parameters of the PV panels used in experiments are given in Table 2.

Table 1. Initial conditions for parameter estimation

At Short Circuit	$[dI/dV]_{sc} = -1/R_{sh,ref}$
At the Maximum Power Point	$[dI/dV]_{sc} = 0$
At Short Circuit Current	$I = I_{sc,ref}, V = 0$
At Open Circuit Voltage	$I = 0, V = V_{oc,ref}$
At the Maximum Power Point	$I = I_{mp,ref}, V = V_{mp,ref}$

Table 2. PV module technical parameters used in analysis

Parameters	Symbol	Values	Units
Solar irradiance	S	1000	W/m ²
Temperature	T	25	°C
Maximum power	P	10	W
Nominal open circuit voltage	V_{ocn}	21.5	V
Voltage at Maximum power	V_{mp}	17.3	V
Nominal short circuit current	I_{scn}	0.71	A
Current at Max. power	I_{mp}	0.57	A

To derive the following equations, PV module initial condition parameters are used in Equation (1) to find short circuit current of the model. Equation (3) represents short circuit current of the PV cell, Equation (4) represents PV cell load current, Equation (5) represents maximum current of the PV cell and Equation (7) represents current to voltage ratio of the model of PV cell respectively. The variables and their symbols used in equations while diode modelling is tabulated in Table 2.

$$I_{sc} = I_L - I_o \left[\exp\left(\frac{I_{sc}R_s}{a_{rf}}\right) - 1 \right] - \frac{I_{sc}R_s}{R_{sh}} \quad (4)$$

$$I_L = I_o \left[\exp\left(\frac{V_{oc}}{a_{rf}}\right) - 1 \right] - \frac{V_{oc}}{R_{sh}} \quad (5)$$

$$I_m = I_L - I_o \left[\exp\left(\frac{V_{mp} + I_{mp}R_s}{a_{rf}}\right) - 1 \right] - \frac{V_{mp} + I_{mp}R_s}{R_{sh}} \quad (6)$$

$$\left[\frac{dI}{dV} \right]_{sc} \cong -\frac{1}{R_{sh}} \quad (7)$$

$$\frac{I_{mp}}{V_{mp}} = \frac{\left(\frac{I_o}{a_{rf}}\right) \exp\left(\frac{V_{mp}I_o + I_{mp}R_s}{a_{rf}}\right) + \frac{1}{R_{sh,rf}}}{1 + \left(\frac{I_oR_s}{a_{rf}}\right) \exp\left(\frac{V_{mp}I_o + I_{mp}R_s}{a_{rf}}\right) + \frac{R_s}{R_{sh}}} \quad (8)$$

The maximum power point efficiency of PV module heavily depends on the cell temperature[18]. PV module efficiency dominantly depends on temperature of module. It is obtained as the following:

$$\eta_{mp} = \frac{I_p V_{mp}}{G_T A_M} = \frac{P_{mp}}{G_T A_M} \quad (9)$$

where, A represents the aperture area and G_T represents irradiance quantity that PV cell exposed, under standard test conditions; 25 °C, $G_T=1000$ W/m². In literature a number of correlations can be found for PV electrical power as a function of cell/module operating temperature and basic environmental variables. Some correlation equations are linear and some of them are nonlinear [19]. As an example, the following multivariable nonlinear regression equation can be shown.

$$P_{mp} = d_1 + d_2 T_c + d_3 (\ln G_T)^m d_4 T_c (\ln G_T)^m \quad (10)$$

Here, $d_j, j=1 \dots 4$ and m are model parameters.

4. Results and Discussion

The experimental set up of PV panel efficiency analysis for two different panel is conducted at Osmaniye Korkutata University roof of Engineering Faculty. Experimental data used in this paper are taken on 12 August 2019 at midday. In the experimental set up two PV module is used.

As previously stated, in this work two PV panels utilized.

First one is unmodified panel, the other one is modified with copper cooling system which is consist of copper plate and pipes. The copper plate is attached back side of PV panel by using high heat transmission coefficient adhesive (Thermal Conductivity: 8.5 W/(m*K)). This type of adhesives can conduct heat from PV panel to copper sheet almost perfectly [20]. Water circulation through the pipes is supplied by natural convection flow. The data from PV panel with and without water cooling is collected simultaneously.

Current-Voltage and Power-Voltage characteristics of PV panels are depicted in figure 6 and figure 7 respectively.

As a result of experimental studies it has been shown that use of a water pipe system is comparatively inexpensive. The results of experiments show that output power of the PV solar system is effectively increased. Real time thermal images of PV panel were also recorded by using maximum ambient temperature. Thermal images shown in Figure 4 and Figure 5, at 13:30 the maximum temperature for a solar panel without and with water cooling 49°C and 40.2°C is observed respectively.

The temperature of the PV panel was varying from 35.2°C to 40°C depending on the sunniest (there were some mist). At the same time, temperature of the water-cooled panel was varying between 34.2°C to 40.2°C respectively. Drop in panel temperature is due to cooling water. The cooling water is supplied from water tank. It has an average temperature of 27 °C. At the beginning, water cooling of PV panel has reduced the surface temperature of the panel below the environment temperature.

As temperature is decreased on PV panel, the efficiency of the panel is enhanced. Power efficiency of water cooled PV panel with respect to non-cooled can be achieved by equation:

$$\eta = \frac{P_{\max \text{ with cooling}} - P_{\max \text{ w/outcooling}}}{P_{\max \text{ w/outcooling}}} \quad (10)$$

Water cooled PV efficiency varies between 5.8 and 6.5 %. During August, midday PV efficiency is measured and its average value is approximately 6.2 %.

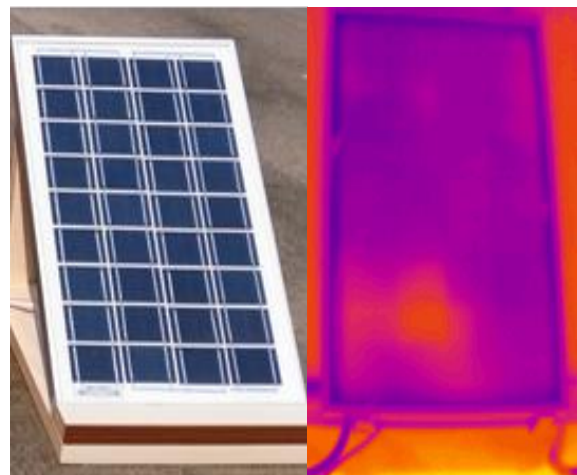


Figure 4. Front view of photovoltaic panel. (left side: Normal image, Right-side: thermal image)

Table 3. Temperature values of water cooled PV panel and environment

	PV with cooling	Environment	PV without cooling
Min temp($^{\circ}$ C)	34.2	35.2	51
Max temp($^{\circ}$ C)	40.2	40	63
Average temp($^{\circ}$ C)	37.5	37	56



Figure 5. Backside, normal and thermal view of developed PV module

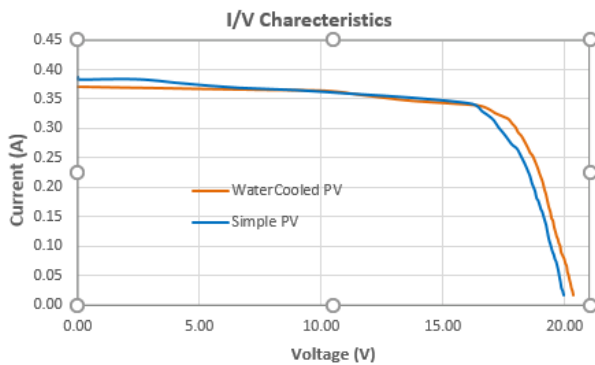


Figure 6. Voltage- Current PV module Characteristics

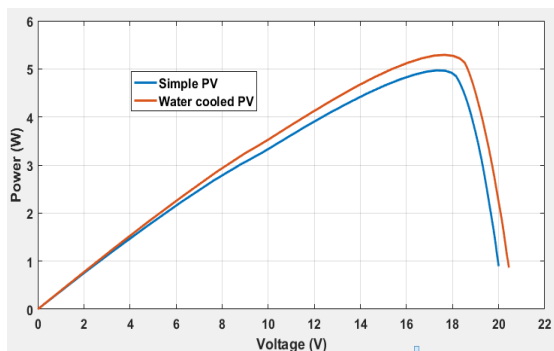


Figure 7. Voltage- Power PV module Characteristics

4. Conclusions

The electrical efficiency of PV panel with water cooling is compared with normal PV panel. The efficiency results of conducted experimental studies are promising. Better results are obtained with respect to previous studies. As a result of cooling, the temperature of panel decreases slightly. Thus,

the PV panel maximum power (P_{mp}) and open circuit voltage (V_{oc}) increase slightly while short circuit current (I_{sc}) decreases slightly. If the overall efficiency of proposed system is considered, water cooled PV system is better than normal PV panel approximately by % 6.2 for experimental test conditions.

The efficiency of PV panel with proposed method is better than most of the results obtained in literature [9, 14].

Declaration

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article. The author(s) also declared that this article is original, was prepared in accordance with international publication and research ethics, and ethical committee permission or any special permission is not required.

Author Contributions

M. Uçman prepared experimental setup. All authors contributed to data collection phase. H. Erol wrote the manuscript. Z. Kesilmiş made proofreading of manuscript.

References

- Bulbul, S., G. Ertuğrul, and F. Arlı, *Investigation of usage potentials of global energy systems*. International Advanced Researches and Engineering, 2018. **2**(1): p. 58-67.
- Fuqiang, W., C. Ziming, T. Jianyu, Y. Yuan, S. Yong, and L. Linhua, *Progress in concentrated solar power technology with parabolic trough collector system: A comprehensive review*. Renewable and Sustainable Energy Reviews, 2017. **79**: p. 1314-1328.
- Xu, J., C.C. Boyd, Z.J. Yu, A.F. Palmstrom, D.J. Witter, B.W. Larson, R.M. France, J. Werner, S.P. Harvey, E.J. Wolf, W. Weigand, S. Manzoor, M.F.A.M. van Hest, J.J. Berry, J.M. Luther, Z.C. Holman, and M.D. McGehee, *Triple-halide wide-band gap perovskites with suppressed phase segregation for efficient tandems*. Science, 2020. **367**(6482): p. 1097-1104.
- Fiducia, T.A.M., B.G. Mendis, K. Li, C.R.M. Grovenor, A.H. Munshi, K. Barth, W.S. Sampath, L.D. Wright, A. Abbas, J.W. Bowers, and J.M. Walls, *Understanding the role of selenium in defect passivation for highly efficient selenium-alloyed cadmium telluride solar cells*. Nature Energy, 2019. **4**(6): p. 504-511.
- Ali, H.M., *Recent advancements in PV cooling and efficiency enhancement integrating phase change materials based systems – A comprehensive review*. Solar Energy, 2020. **197**: p. 163-198.
- Shafique, M., X. Luo, and J. Zuo, *Photovoltaic-green roofs: A review of benefits, limitations, and trends*. Solar Energy, 2020. **202**: p. 485-497.
- Kumar, P., A.K. Shukla, K. Sudhakar, and R. Mamat, *Experimental exergy analysis of water-cooled PV module*. International Journal of Exergy, 2017. **23**(3): p. 197-209.
- Bianchini, A., A. Guzzini, M. Pellegrini, and C. Saccani, *Photovoltaic/thermal (PV/T) solar system: Experimental measurements, performance analysis and economic assessment*. Renewable Energy, 2017. **111**: p. 543-555.

9. Erol, H., M. Uçman, and Z. Kesilmiş, *The effect of fan cooling on Photovoltaic Efficiency of PV panels in Osmaniye Environment*. *Majlesi Journal of Mechatronic Systems*, 2017. **6**(3): p. 29-33.
10. Bahaidarah, H., A. Subhan, P. Gandhidasan, and S. Rehman, *Performance evaluation of a PV (photovoltaic) module by back surface water cooling for hot climatic conditions*. *Energy*, 2013. **59**: p. 445-453.
11. Clean energy reviews. [cited 2020 29 June]; Available from: <https://www.cleanenergyreviews.info/blog/most-efficient-solar-panels>.
12. Huang, G., S.R. Curt, K. Wang, and C.N. Markides, *Challenges and opportunities for nanomaterials in spectral splitting for high-performance hybrid solar photovoltaic-thermal applications: A review*. *Nano Materials Science*, 2020. **2**(3): p. 183-203.
13. Ren, X., J. Li, M. Hu, G. Pei, D. Jiao, X. Zhao, and J. Ji, *Feasibility of an innovative amorphous silicon photovoltaic/thermal system for medium temperature applications*. *Applied Energy*, 2019. **252**: p. 113427.
14. Matias, C.A., L.M. Santos, A.J. Alves, and W.P. Calixto, *Electrical performance evaluation of PV panel through water cooling technique*. 2016 IEEE 16th International Conference on Environment and Electrical Engineering (Eeeic), 2016. Florence, Italy.
15. Matias, C.A., L.M. Santos, A.J. Alves, and W.P. Calixto, *Increasing photovoltaic panel power through water cooling technique*. *Transactions on Environment and Electrical Engineering*, 2017. **2**(1): p. 60-66.
16. Hui, G., Q. Da-yong, W. Yue-hu, Z. Yu-ming, and Z. Yi-men. *The intermediate semiconductor layer for the ohmic contact to silicon carbide by Germanium implantation*. in *Junction Technology*, 2008. *IWJT '08. Extended Abstracts - 2008 8th International workshop on*. 2008.
17. Gaikwad, K. and S. Lokhande. *Novel maximum power point tracking (MPPT) algorithm for solar tree application*. in *2015 International Conference on Energy Systems and Applications*. 2015.
18. Kesilmiş, Z., H. Erol, and M. Uçman, *Power optimization in partially shaded photovoltaic systems*. *Tehnički glasnik*, 2018. **12**: p. 34-38.
19. Rosell, J.I. and M. Ibáñez, *Modelling power output in photovoltaic modules for outdoor operating conditions*. *Energy Conversion and Management*, 2006. **47**(15): p. 2424-2430.
20. Uçman, M., *Fotovoltaik termal (FV/T) ve bina entegreli FV/T (BIPV/T) sistemlerde akışkan kullanımının elektriksel karakteristik üzerine etkisinin incelenmesi* in *Electrical and Electronics Engineering*. 2017, Osmaniye Korkut Ata University: Turkey. p. 111.



Research Article

Silicon dioxide thin films prepared by spin coating for the application of solar cells

İmran Kanmaz^{a,*}  and Abdullah Üzüm^{a,b} 

^aKaradeniz Technical University, Department of Renewable Energy Resources/Technologies, Trabzon, Turkey

^bKaradeniz Technical University, Department of Electrical and Electronics Engineering, Trabzon, Turkey

ARTICLE INFO

Article history:

Received 24 August 2020

Revised 28 October 2020

Accepted 06 November 2020

Keywords:

Afors-het

Antireflection Coating

c-Si Solar cells

SiO₂

Thin Films

ABSTRACT

In this study, Silicon Dioxide (SiO₂) thin films processed by the spin coating method was studied with prepared solutions. Antireflection coating effect of deposited SiO₂ thin films on crystalline silicon substrates was analyzed after optimizing the solution, deposition, and thermal treatment processes. The effect of ethanol dilution of the solution was investigated as well. Spectrophotometer reflectance measurements, Scanning Electron Microscopy (SEM) measurements and Afors-het based simulations were carried out. For the prepared solution based SiO₂ thin films, the annealing temperature of 950 °C for 7 min in the air was determined as optimum. The minimum surface reflectance of SiO₂ coated silicon surface could be reduced below 10% depending on the applied process. Based on the silicon solar cell device simulations, it was revealed that efficiency of a solar cell could be improved 4.23% more thanks to the antireflection coating effect.

© 2021, Advanced Researches and Engineering Journal (IAREJ) and the Author(s).

1. Introduction

The reflection of the light coming on the surface of the solar cells is an important factor affecting the efficiency [1]. Anti-reflective layers (ARLs) are playing an important role in minimizing reflection losses on the surface of solar cells. By applying one or more ARL layers to photovoltaic cells [2] or to other optoelectronic devices, the reflection of incoming light can be reduced, thus the performance of the device can be improved [3, 4]. However, high production and raw material expenses increase the cost per unit watt. Energy costs can be reduced by reducing costs in production methods. As an effective thin film coating approach, the ARLs can be easily prepared by sol-gel based coating processes such as dipping, spinning, meniscus, and spray pyrolysis to obtain quality optical thin films [5-8]. Compared to physical and chemical steam methods, the sol-gel method has many advantages such as being cheaper, easily adaptable to industry scale and mass production, working under normal atmospheric conditions, high purity, ultra-homogeneity, and working at the desired

molarity and temperature [9-11]. For anti-reflection coatings, highly transparent materials such as SiO₂, Si₃N₄, TiO₂, Al₂O₃, Ta₂O₅, and SiO₂-TiO₂ are used for c-Si solar cells [12].

High-quality anti-reflection film coatings are generally produced by methods that require vacuum, such as thermal evaporation, atomic layer deposition, chemical vapor deposition, or plasma-enhanced chemical vapor deposition method, etc. [13]. However, these methods are quite costly compared to the sol-gel method. In addition, most of these methods are performed with atmosphere control. On the other hand, the sol-gel method including spin coating is generally carried out in ambient air. The solution needs to be prepared properly to meet the aim of application which will enhance antireflection properties in current study.

In this study, SiO₂ thin films created by the spin coating method and their antireflection properties were investigated. The used solution shows differences in terms of content ratios and experimental stages compared to the literature. In addition to the explained preparation and optimization process of the SiO₂ solution, the effect of

* Corresponding author. Tel.: +90 462 377 4279

E-mail addresses: imrankanmaz@ktu.edu.tr (I. Kanmaz), auzum@ktu.edu.tr (A. Üzüm)

ORCID: 0000-0001-8827-1590 (I. Kanmaz), 0000-0001-5324-8892 (A. Üzüm)

DOI: 10.35860/iarej.784328

This article is licensed under the CC BY-NC 4.0 International License (<https://creativecommons.org/licenses/by-nc/4.0/>).

ethanol dilution of the solution was investigated as well. Moreover, the effect of ARL on c-Si solar cells obtained by SiO₂ thin films was observed using Afors-het simulations.

2. Experimental

The initial SiO₂ solution was prepared with volume ratio of 3: 2: 4; TEOS (C₈H₂₀O₄Si - %99), Ethyl alcohol and water were mixed with a magnetic stirrer for half an hour at room temperature. Then, 150 µL of HCl was added to clear the solution. Subsequently, 8 mL of ethyl alcohol was added to the prepared solution to dilute it and it was kept at room temperature for 24 hours (would be called "SiO₂-Sol" hereafter). Polished p-type CZ-Si with 725 µm thickness and 1.5 cm × 1.7 cm dimensions was used as a substrate. Substrates were cleaned with ethanol in an ultrasonic bath for 15 min and then air-dried. Then, they were further cleaned with purified water in ultrasonic bath for 15 min and finally the drying step in air was followed.

Coating was carried out using the spin coating process. SiO₂ thin films were coated on the Si substrate at 4000 rpm for 30 s. Annealing temperature, annealing time, and thin film thickness optimizations were carried out in order to achieve the minimum reflectance value for the thin film coated on the c-Si surface. To adjust the desired film thickness of the optimized solution, different amounts of solution (100 µL, 50 µL, 35 µL and 25 µL) were dropped onto the substrate surface and the coating was completed. Subsequently, samples were prepared to adjust the annealing temperature in atmosphere for 10 min from 550 °C to 1000 °C. To adjust the annealing time, annealing was performed at different temperatures for periods from 1 min to 30 min. Moreover, in order to observe the effect of further dilution of the initial SiO₂-Sol, four different diluted solutions of Sol:Ethanol (1:1, 1:2, 1:5, 1:10 in volume) were prepared. Reflectance behaviors of the films coated by further diluted SiO₂-Sols were compared to that of the films coated by undiluted SiO₂-Sol. Optical and morphological characterizations of thin films were performed with the Spectrophotometer reflection measurements and Scanning Electron Microscopy (SEM) measurements. Finally, simulations of c-Si solar cells were performed using Afors-het software.

3. Result and Discussion

To examine the characteristic properties of the coated SiO₂ thin films with varying thicknesses, thin films were formed by the spin coating method by dropping the same molarity solution in different amounts on the substrate. Depending on the amount of solution used on the substrate, the corresponding reflectance spectra of coated Si surfaces are shown in figure 1. In addition, the minimum reflectance and average reflectance values of this graph are shown in Table 1 in the range of 350 nm to 1000 nm. It

was observed that there was a shift in the minimum point of the reflectance spectra with the amount of solution dripped onto the substrate. The shift can be attributed to the increase of the formed film thickness [14, 15]. The minimum reflectance value increased from 9.60% to 11.50% as the amount of solution increased Table 1.

In addition, it was observed that the average reflectance value varied from 19.02% to 19.62% depending on the increasing amount of solution. 35 µL was set as the optimum amount of drop for the rest of the experiments because it was in the appropriate wavelength range and provided lower average reflectance (19.02%) than that of the other samples.

By using thin films created under the same conditions with the same experimental parameters, annealing temperature optimization was performed between 550 °C and 1000 °C for 10 min. The reflectance spectra of the samples prepared depending on the temperature were shown in figure 2. In addition, Table 2 presents the minimum and the average reflectance values for these samples in the wavelength ranging from 350 nm to 1000 nm.

The minimum reflectance value increased from 10.30% to 11.50% when the temperature increased from 550 °C to 750 °C; however, after 750 °C, it decreased again to 9.50% when the annealing temperature was 1000 °C (Table 2). Due to the minimum and average reflectance values as well as the appropriate wavelength range, 950 °C was considered as the optimum temperature.

Table 1. Minimum and average reflection values depending on the amount of SiO₂-Sol coated on the Si surface

Amount (µL)	Average Ref. (350-1000 nm)	Min. Ref (%)
100	19.26	11.50 (555nm)
50	19.62	11.00 (585nm)
35	19.02	10.10 (615nm)
25	19.20	9.60 (675nm)

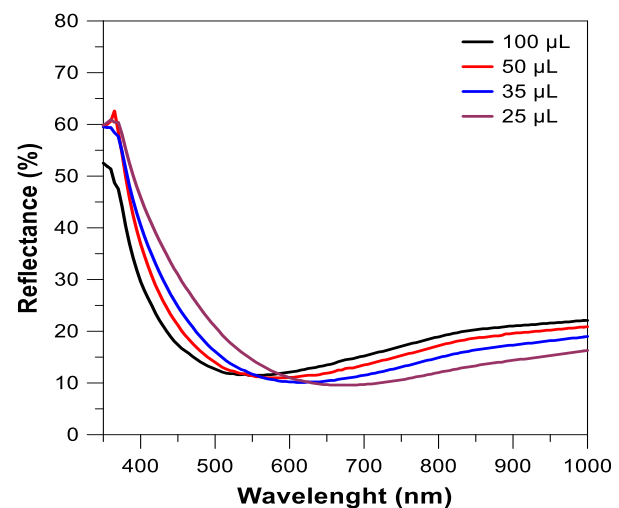


Figure 1. Reflectance spectra of the Si surface depending on the amount of coated SiO₂-solution

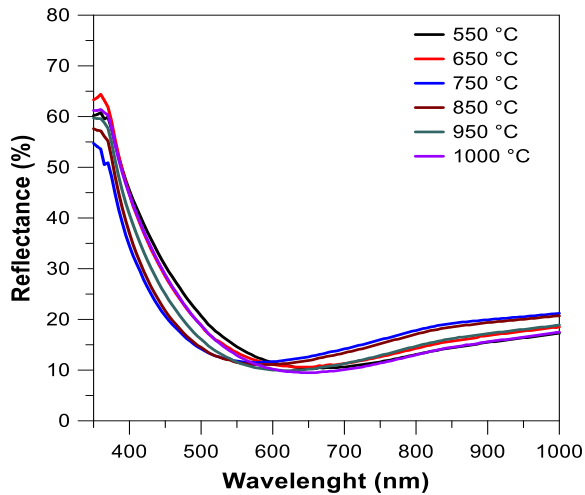


Figure 2. Reflectance of SiO₂-coated-Si surface depending on different annealing temperatures

Table 2. Minimum and average reflection values depending on annealing temperature

Temperature (°C)	Average Ref. (350-1000nm)	Min. Ref (%)
550	19.74	10.30 (660nm)
650	19.98	10.60 (645nm)
750	19.47	11.50 (580nm)
850	19.48	11.10 (590nm)
950	18.93	9.90 (620nm)
1000	19.09	9.50 (650nm)

The prepared samples were annealed in the air atmosphere at 950 °C from 1 min to 30 min in order to observe the effect of the annealing time. Reflectance spectra by annealing time are shown in figure 3.

As the annealing time increased from 1 min to 30 min, minimum reflectance values decreased continuously from 13.50% to 9.40%. In contrast to the minimum reflectance value, although the average reflectance values decreased from 20.61% to 19.26% when the annealing time was increased from 1 min to 10 min, it was increased up to 20.06% at 30 min.

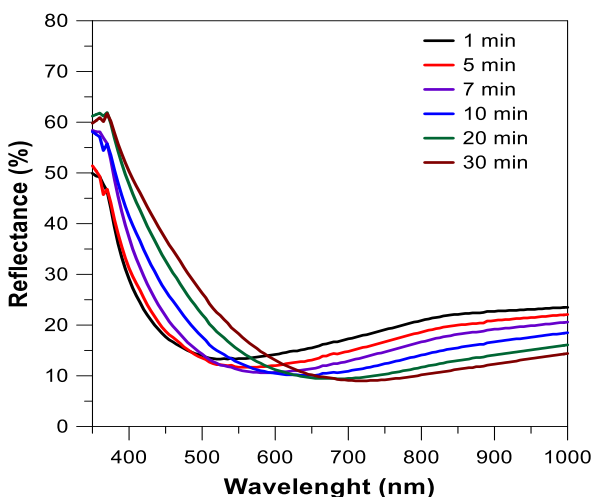


Figure 3. Reflectance of SiO₂-coated-Si surface depending on different annealing times

Table 3. Minimum and average reflection values depending on annealing time

Time (min)	Average Ref. (350-1000 nm)	Min. Ref (%)
1	20.61	13.30 (535nm)
5	19.34	11.70 (565nm)
7	19.26	10.60 (590nm)
10	18.99	10.10 (645nm)
20	19.54	9.40 (690nm)
30	20.06	9.00 (720nm)

Since the dilution of the sol-gel prepared solutions with ethanol was reported in the literature [16, 17], attention was given to the effect of ethanol dilution on the reflection properties of the resulted thin films. The SiO₂-Sol is the optimized solution in terms of smooth and uniform coatability. The effect of further dilution of the SiO₂-Sol with ethanol (Sol:Ethanol in volume, 1:1, 1:2, 1:5, 1:10) on the antireflection behavior can be seen in figure 4. Reflectance of the c-Si surface was increased by adding additional ethanol into the solution comparing with that of the SiO₂-Sol. Moreover, the minimum reflection point was apart from the desired visible region. Therefore, no improvements could be observed at the end of further dilution of the SiO₂-Sol.

Based on the results, annealing temperature of 950 °C for 7 min in air was determined to be the optimum conditions for the prepared SiO₂ thin films. As seen in figure 4, the average reflectance value decreased from approximately 40% (bare c-Si surface) to 19.26% with the coating of SiO₂ thin film on the c-Si surface.

Figure 5 shows morphology of the optimum SiO₂ thin film formed on c-Si substrate with coating speed of 4000 rpm and annealed at 950 °C for 7 minutes. It is observed that the nanostructured SiO₂ film has been formed and possesses a porous structure.

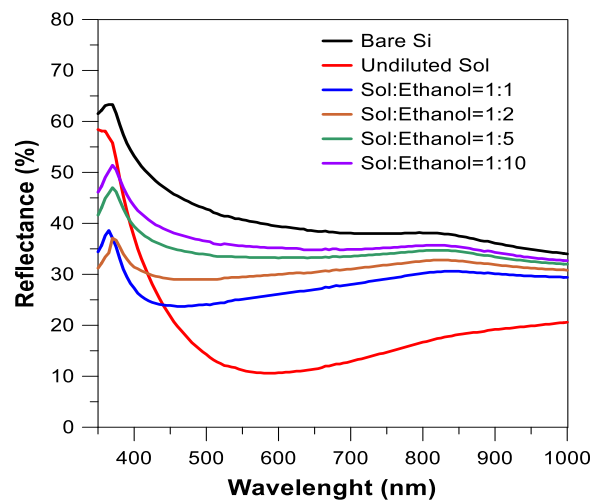


Figure 4. Reflectance spectra of the c-Si surface depending on the further dilution of SiO₂-solution

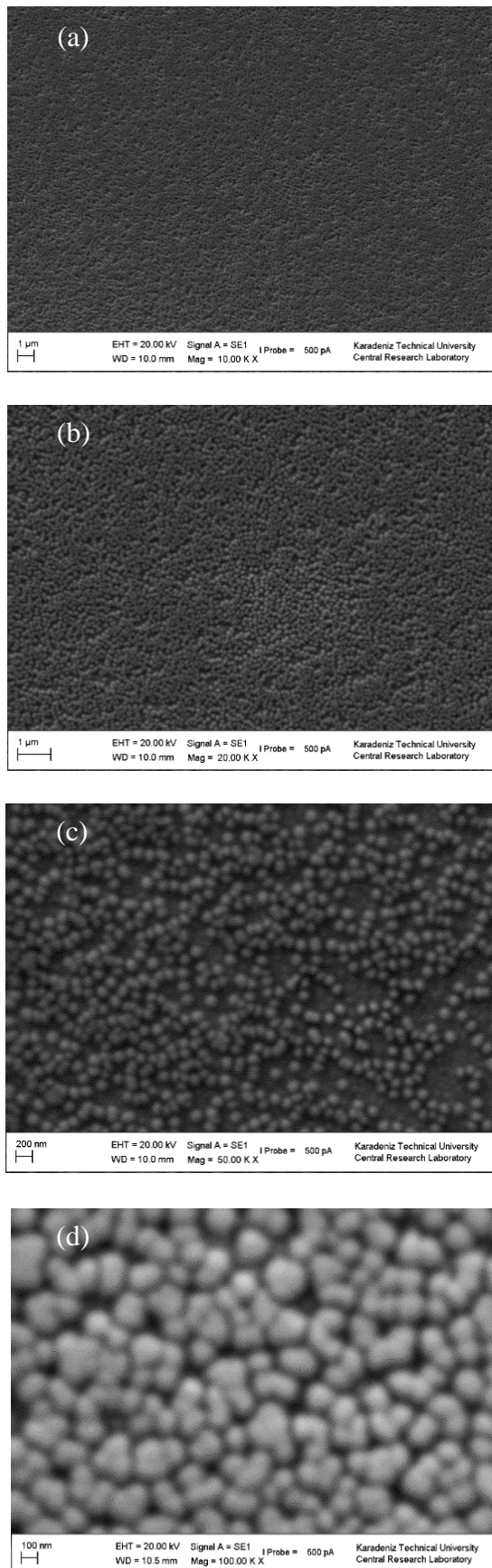


Figure 5. SEM micrographs of SiO₂ thin films a) $\times 10k$, b) $\times 20k$, c) $\times 50k$, d) $\times 100k$

The average reflectance values obtained from the optimum SiO₂/c-Si structures were applied to the c-Si solar cell as ARL by using Afors-het simulation [18]. For c-Si

solar cell, some initial parameters including emitter thickness, emitter doping concentration, base thickness, base doping concentration, back surface field (BSF) thickness, BSF doping concentration, series resistance, and parallel resistance were determined in Afors-het software (Table 4). I-V graph for c-Si solar cell with or without SiO₂ coating as an anti-reflection coating is shown in figure 6 with the characteristic parameters of solar cells given in Table 5.

While the power conversion efficiency (PCE) of the solar cell without ARL coating was determined as 11.68%, the efficiency of the solar cell with reduced reflection was obtained as 15.91%. A 4% increase in efficiency was achieved as a result of an increase in the short-circuit current density (J_{sc}) from 20.87 to 28.33 mA/cm² and an increase in the open-circuit voltage (V_{oc}) from 685.1 to 693.40 mV due to a decrease in reflection on the silicon surface. Similar impacts of the increase of J_{sc} and V_{oc} on PCE has also been reported elsewhere practically [19, 20], which confirms the importance and great impact of ARLs on the performance of solar cells.

Table 4. Some initial parameters set in Afors-het for c-Si solar cell

Emitter thickness (μm)	0.3
Emitter doping concentration (cm^{-3})	1.0×10^{20}
Base Thickness (μm)	200
Base doping concentration (cm^{-3})	1.5×10^{16}
BSF doping concentration (cm^{-3})	1.0×10^{19}
BSF thickness (μm)	7
Series resistance ($\Omega \text{ cm}^2$)	0.8
Parallel resistance ($\Omega \text{ cm}^2$)	10000

Table 5. Characteristic parameters of SiO₂ coated and uncoated solar cell as ARL

	J_{sc} (mA/cm ²)	V_{oc} (mV)	FF (%)	Eff (%)
non-ARL	20.87	685.10	81.64	11.68
with-ARL	28.33	693.40	81.01	15.91

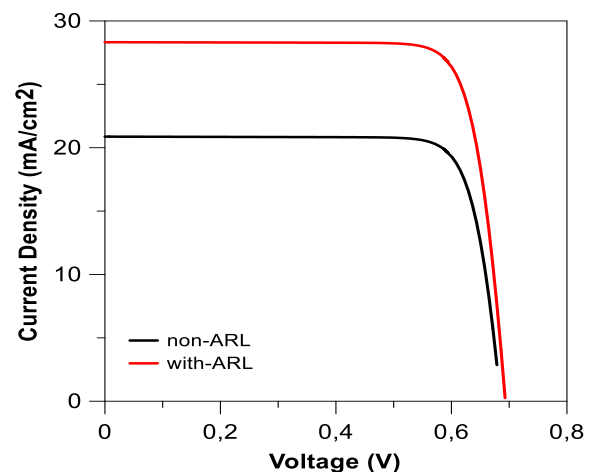


Figure 6. I-V graph of an uncoated and SiO₂ coated solar cell

4. Conclusion

In this study, SiO₂ based spin coating processed ARL thin films were introduced. ARL effect of the prepared SiO₂ films were provided with the recipe of the optimum solution and the optimum processing conditions. Average reflection (350 – 1000 nm) of c-Si surface could be reduced from 40.5% to 19.26%; the minimum of it could be <10% depending on the process. Simulations related to the effect of ARL on c-Si solar cell device were also shown by increasing the PCE of a solar cell from 11.68% to 15.91%. These results can contribute to the development of solar cells through solution-based, non-vacuum, and simple spin-coating processing.

Declaration

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article. The authors also declared that this article is original, was prepared in accordance with international publication and research ethics, and ethical committee permission or any special permission is not required.

Author Contributions

I. Kanmaz and A. Üzüm conceived the study together. I. Kanmaz conducted the experiments, analyzed the data and drafted the paper. A. Üzüm supervised the research, contributed the data analyze and revised / finalized the paper.

Acknowledgement

This research was supported by TUBITAK (The Scientific and Technological Research Council of Turkey) under the project number of 119F063.

References

- Zhan, F., Z. Li, X. Shen, H. He, and J. Zeng, *Design multilayer antireflection coatings for terrestrial solar cells*. The Scientific World Journal, 2014. **265351**.
- Lien, S.-Y., D.-S. Wu, W.-C. Yeh, and J.-C. Liu, *Tri-layer antireflection coatings (SiO₂/SiO₂-TiO₂/TiO₂) for silicon solar cells using a sol-gel technique*. Solar Energy Materials and Solar Cells, 2006. **90**(16): p. 2710-2719.
- Bouhafs, D., A. Moussi, A. Chikouche, and J. Ruiz, *Design and simulation of antireflection coating systems for optoelectronic devices: Application to silicon solar cells*. Solar Energy Materials and Solar Cells, 1998. **52**(1-2): p. 79-93.
- Ko, J., D. Gong, K. Pillai, K.-S. Lee, M. Ju, P. Choi, K.-R. Kim, J. Yi, and B. Choi, *Double layer SiN_x: H films for passivation and anti-reflection coating of c-Si solar cells*. Thin solid films, 2011. **519**(20): p. 6887-6891.
- Bacal, D.M., N.N. Lal, A.N. Jumabekov, Q. Hou, Y. Hu, J. Lu, A.S. Chesman, and U. Bach, *Solution-processed antireflective coating for back-contact perovskite solar cells*. Optics Express, 2020. **28**(9): p. 12650-12660.
- Fathima, M.I. and K.J. Wilson. *Antireflection coating application of zinc sulfide thin films by nebulizer spray pyrolysis technique*. in *AIP Conference Proceedings*. 2019. AIP Publishing LLC.
- Reddy, K.C.S., C. Chingakham, B. Gupta, M.S. Prasad, S. Atchuta, and S. Sakthivel, *Single compound in-situ synthesis of core-shell CaF₂ nanoparticles based broad band antireflective coatings for solar energy conversion*. Solar Energy, 2019. **190**: p. 119-125.
- Saranam, V.R., C.-H. Chang, and B.K. Paul, *A Foam-Core Meniscus Coating Process for Retrofit Anti-Reflective Coatings*. Journal of Micro and Nano-Manufacturing, 2019. **7**(3).
- Hinczewski, D.S., M. Hinczewski, F. Tepehan, and G. Tepehan, *Optical filters from SiO₂ and TiO₂ multi-layers using sol-gel spin coating method*. Solar energy materials and solar cells, 2005. **87**(1-4): p. 181-196.
- Jeon, H.-J., S.-C. Yi, and S.-G. Oh, *Preparation and antibacterial effects of Ag-SiO₂ thin films by sol-gel method*. Biomaterials, 2003. **24**(27): p. 4921-4928.
- Zheng, C., F. Nie, Y. Zheng, Y. Cheng, S. Wei, L. Ruan, and R. Valiev, *Enhanced corrosion resistance and cellular behavior of ultrafine-grained biomedical NiTi alloy with a novel SrO-SiO₂-TiO₂ sol-gel coating*. Applied Surface Science, 2011. **257**(13): p. 5913-5918.
- Wuu, D.-S., C.-C. Lin, C.-N. Chen, H.-H. Lee, and J.-J. Huang, *Properties of double-layer Al₂O₃/TiO₂ antireflection coatings by liquid phase deposition*. Thin Solid Films, 2015. **584**: p. 248-252.
- Liang, Z., W. Li, B. Dong, Y. Sun, H. Tang, L. Zhao, and S. Wang, *Double-function SiO₂-DMS coating with antireflection and superhydrophobic surface*. Chemical Physics Letters, 2019. **716**: p. 211-214.
- Cho, E.-C., J. Xia, A.G. Aberle, and M.A. Green, *Antireflection and surface passivation behaviour of SiO₂/Si/SiO₂ quantum wells on silicon*. Solar energy materials and solar cells, 2002. **74**(1-4): p. 147-154.
- Miao, L., L.F. Su, S. Tanemura, C.A. Fisher, L.L. Zhao, Q. Liang, and G. Xu, *Cost-effective nanoporous SiO₂-TiO₂ coatings on glass substrates with antireflective and self-cleaning properties*. Applied energy, 2013. **112**: p. 1198-1205.
- Wang, J., J. Ge, H. Hou, M. Wang, G. Liu, G. Qiao, and Y. Wang, *Design and sol-gel preparation of SiO₂/TiO₂ and SiO₂/SnO₂/SiO₂-SnO₂ multilayer antireflective coatings*. Applied Surface Science, 2017. **422**: p. 970-974.
- Guglielmi, M., A. Martucci, R. Almeida, H. Vasconcelos, E. Yeatman, E. Dawnay, and M. Fardad, *Spinning deposition of silica and silica-titania optical coatings: A round robin test*. Journal of materials research, 1998. **13**(3): p. 731-738.
- Varache, R., C. Leendertz, M. Gueunier-Farret, J. Haschke, D. Muñoz, and L. Korte, *Investigation of selective junctions using a newly developed tunnel current model for solar cell applications*. Solar Energy Materials and Solar Cells, 2015. **141**: p. 14-23.
- Ali, K., S.A. Khan, and M.M. Jafri, *Effect of double layer (SiO₂/TiO₂) anti-reflective coating on silicon solar cells*. Int. J. Electrochem. Sci, 2014. **9**(12): p. 7865-7874.
- Salman, K.A., K. Omar, and Z. Hassan, *Effective conversion efficiency enhancement of solar cell using ZnO/PS antireflection coating layers*. Solar Energy, 2012. **86**(1): p. 541-547.

**Research Article****Selective harmonic elimination in multi-level inverters by using neural networks**Özkan Akin ^{a,*} , İbrahim Özer ^a  and Halil Ünlü ^a ^aEge University, Department of Electrical and Electronics Engineering, Izmir 35100, Turkey

ARTICLE INFO

Article history:

Received 19 August 2020

Revised 20 February 2021

Accepted 04 March 2021

*Keywords:*Artificial neural networks
Multilevel inverter
Newton-Raphson method
Selective harmonic elimination

ABSTRACT

In this study, it is aimed to eliminate the harmonics selected by the selective harmonic elimination (SHE) method in a 7-level cascade multilevel inverter using artificial neural networks (ANNs). A control algorithm has been developed in which the 3rd and 5th harmonics or 5th and 7th harmonics can be eliminated according to the selection while adjusting the output voltage amplitude of the inverter. The required switching angles for SHE are calculated in real time using ANN. These angles were first obtained offline training of ANN using Newton-Raphson method. ANN was trained in MATLAB® environment according to the obtained data. The resulting ANN algorithm and practical implementation using the STM32F429 ARM microcontroller® and inverter switching was provided. Experimental results of the system with RL load were tested.

© 2021, Advanced Researches and Engineering Journal (IAREJ) and the Author(s).

1. Introduction

Inverters are electrical power converters that convert DC voltage to AC voltage and are used in many areas such as motor drivers and renewable energy grid integration. In this area, multi-level inverters stand out more than traditional two-level inverters due to their advantages such as low total harmonic distortion (THD) value, better power quality, higher amplitude of the fundamental component and better electromagnetic compatibility [1]. Multilevel inverters can be regarded as voltage synthesizers where the desired output voltage is generated by the independent smaller DC voltage. Increasing the number of DC voltage sources causes the converter output voltage to reach an almost sinusoidal waveform.

Multilevel topologies are classified into three major categories as diode-clamped, flying capacitor, and cascading multilevel structures. Stepped multi-level structures are preferred more with their many advantages over other topologies. The most important advantage of cascading multilevel topology over diode-clamped and flying capacitors is its simplicity. While diode-clamped

and flying capacitor configurations require more capacitors and diodes, cascading topology requires more isolated DA sources. Hence the complexity of the control strategy for controlling the voltage across each capacitor is the main drawback in flying capacitors and diode clamp converters. Requiring more insulated voltage sources may be disadvantageous depending on the type of application. However, they provide advantages in photovoltaic applications or applications in battery powered systems due to the independent nature of the sources [2].

There are many switching methods that are divided into two groups according to the switching frequency of multi-level inverters. Classical carrier-based Sinusoidal Pulse Width Modulation (SPWM) and Space Vector Modulation (SVM) are some of the high frequency methods. Selective Harmonic Elimination (SHE) method is available as the low frequency method. SHE method is preferred over high frequency methods because it provides better harmonic profile and lower switching loss [3]. The main problem in the SHE method is to find the desired output voltage and the angles required to eliminate the selected harmonics. The challenge here is

* Corresponding author. Tel.: +90-232-311-5248; Fax: +90-232-388-6024.

E-mail addresses: ozkan.akin@ege.edu.tr (Ö. Akin), ibrahimozzer@gmail.com (İ. Özer), hunlu9507@gmail.com (H. Ünlü)

ORCID: 0000-0002-2214-3374 (Ö. Akin), 0000-0001-6335-4576 (İ. Özer), 0000-0001-9473-0618 (H. Ünlü)

DOI: 10.35860/iarej.782492

This article is licensed under the CC BY-NC 4.0 International License (<https://creativecommons.org/licenses/by-nc/4.0/>).

the complexity of solving higher-order nonlinear equations that need to be solved to calculate switching angles. There are many algorithms such as Newton-Raphson (NR), Sequential Quadratic Programming (SQP), Result Theory, Homotopy Algorithm, Genetic algorithm (GA), PSO and Harmony Search Algorithm (HSA) to find these angles [4]. However, solving the equations with these algorithms is very time consuming and difficult to implement in real-time for today's microprocessor speeds. Therefore, the necessary angles in applications are obtained offline by using these algorithms and a lookup table is created [5, 6]. While this method is applicable for basic applications, when the flexibility of the application increases, such a lookup table will require an amount of memory and fast processing that can easily exceed the processor capacity. The idea of using artificial neural networks (ANN) to overcome this problem has been proposed in recent years.

Artificial neural networks are a computing technology inspired by the information processing technique of the human brain. ANN has artificial nerve cells similar to human nerve cells. These cells determine at what rate the data received with the elements they have will be transmitted to the next cell. Therefore, cell elements should be updated until the desired input-output value is reached for ANN training.

ANN has a very fast calculation capability as a result of a good training. This feature has enabled it to find a place in power electronics applications where processing speed is very critical [7]. ANN can be trained according to the needs of each study with many training methods [8]. In SHE applications, the ANN must calculate the switching angles in real time. Therefore, ANN should be trained on an angle table obtained by old methods.

There are some SHE implementations made using ANN. The most common studies are studies that provide constant output voltage at variable input voltages, which are more suitable especially for solar systems [9, 10]. In these studies, the equation coefficients to be solved are constantly changing due to variable cell voltages. Therefore, angle samples taken for ANN training can be quite large. Another study is applications that give variable output voltage for constant input voltage, made for applications that require variable output voltage such as motor control [11, 12]. Since the equation coefficients are constant in these studies, ANN training gives better results than other studies.

In the literature, studies in general have been carried out for a single purpose, especially connecting the output voltage to the network. In this study, an inverter design that can be used for many different applications is realized. For a 7-level inverter, up to two harmonics can be eliminated simultaneously. Since the multiples of three (triple) harmonics are not important in systems connected

to the grid, it is sufficient to eliminate the 5th and 7th harmonics. However, 3rd harmonic is a major disadvantage for single phase motor drive applications. For this, the options to eliminate 3rd and 5th or 5th and 7th harmonics, which can be selected depending on the application, are presented to the user with an interface. In addition, fixed and variable frequency options are offered by considering the V/f control, which is frequently used in motor drive applications. In this study, a 7-level inverter design, which includes all these features, has been implemented. Thus, unlike other studies, a different perspective has been brought to this field with an inverter design that can be used for many purposes.

2. System Introduction and Working Principles

In the study, two different harmonic selection, in which 3rd and 5th or 5th and 7th harmonics are eliminated, two different frequency options for constant and V/f control, and a control system that includes modulation index change options for variable output voltage, providing the appropriate output waveform. 7-level level inverter control has been implemented. The cascaded inverter is obtained by serially connecting the H-Bridge circuits powered by independent DC sources as shown in Figure 1. This inverter output is provided by switching angle control for each bridge circuit as seen in Figure 2.

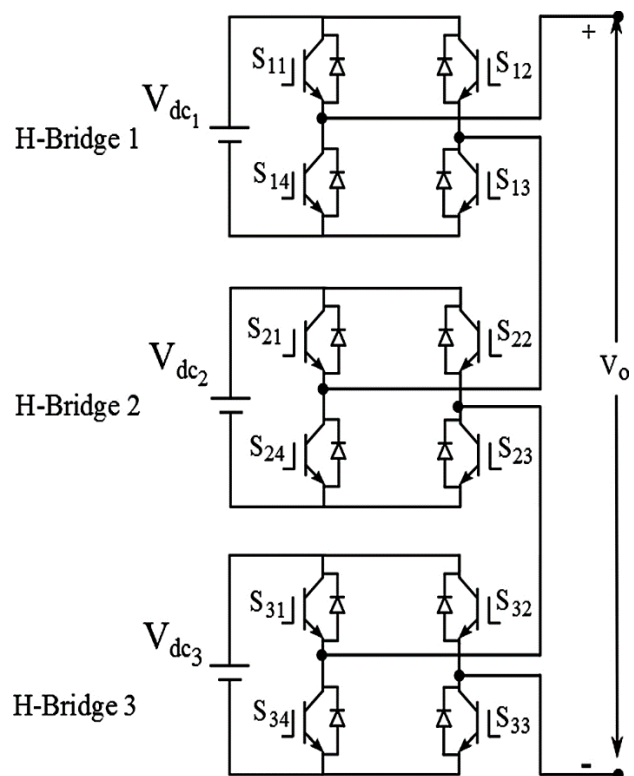


Figure 1. Single phase cascade h-bridge multilevel inverter

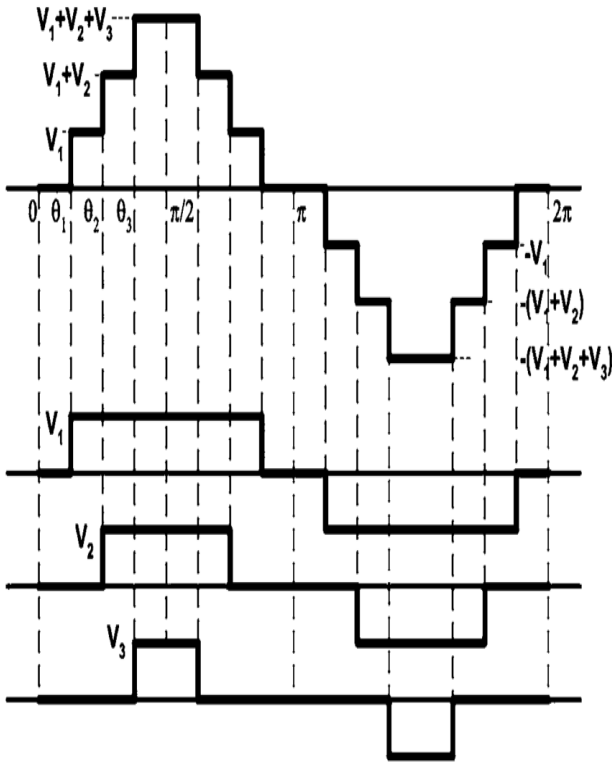


Figure 2. Staircase voltage waveform for single-phase multilevel inverter.

Fourier series expansion for the output waveform in Figure 2 is as in Equation (1).

$$V_{out}(\omega t) = \sum_{n=1,3,5...}^{\infty} b_n \sin(n\omega t) \quad (1)$$

b_n is given by:

$$b_n = \sum_{n=1,3,5...}^{2N-1} (V_{dc1} \cos(n\alpha_1) + V_{dc2} \cos(n\alpha_2) + \dots + V_{dcL-1} \cos(n\alpha_{L-1}) + V_{dcL} \cos(n\alpha_L) + \dots) \quad (2)$$

For equal and constant source, b_n is given by;

$$b_n = \sum_{n=1,3,5...}^{2N-1} V_{dc} (\cos(n\alpha_1) + \cos(n\alpha_2) + \dots + \cos(n\alpha_{L-1}) + \cos(n\alpha_L) + \dots) \quad (3)$$

Where,

$$V_{dc} = V_{dc1} = V_{dc2} = \dots = V_{dcL}$$

L = number of dc sources for each full H-bridge inverter cell,

N = number of switching angles,

$n=1, 3, 5, \dots$ odd harmonics $(2N-1)$.

Harmonic equations including the basic voltage equation are derived from Equation (3). In the SHE method, as seen in Equation (4), the desired value for the basic component is assigned and the harmonic equation to be eliminated is set to zero. The desired output voltage and harmonic elimination can be achieved with the switching angles that give the solution of these equations.

$$\begin{aligned} \cos(\alpha_1) + \cos(\alpha_2) + \dots \cos(\alpha_L) &= m \\ \cos(3\alpha_1) + \cos(3\alpha_2) + \dots \cos(3\alpha_L) &= 0 \\ \cos(5\alpha_1) + \cos(5\alpha_2) + \dots \cos(5\alpha_L) &= 0 \\ \cos(7\alpha_1) + \cos(7\alpha_2) + \dots \cos(7\alpha_L) &= 0 \\ \cos(n\alpha_1) + \cos(n\alpha_2) + \dots \cos(n\alpha_L) &= 0 \end{aligned} \quad (4)$$

Where:

Modulation index, $m = V_1 / (4V_{dc} / \pi)$, L is the number of dc sources.

The main challenge is solving these nonlinear equations. An iterative algorithm such as the Newton-Raphson method is traditionally used to solve these equivalents. However, since the angle calculation is not possible in the output voltage frequency with this algorithm, it cannot be performed in real time. As stated before, angle tables calculated offline are useless because they can easily exceed the microcontroller memory. This problem has been solved by using ANN. ANN has the ability to do the same process faster by being trained with the input-outputs of time consuming transactions. Figure 3 shows the structure of a single neuron in ANN. For a single neuron, the processing charge has the power of very fast processing since it consists of passing the input value multiplied by the weight value through a transfer function and transmitting it to the next neuron as in Equation (5).

$$o_j = \varphi\left(\sum_{k=1}^n w_k x_k + \theta_j\right) \quad (5)$$

The training of ANN is based on updating the weights and bias values that make up its structure until it gives the most reliable result.

In the study, it is aimed to design a 7-level multilevel inverter. The following options are provided for the inverter output waveform; variable output voltage in the range of 40% -90% modulation index ($m = 1.2-2.7$) for the fundamental component, elimination of 3rd-5th harmonics or 5th-7th harmonics in this modulation index range, fixed output frequency according to selection or variable output frequency in V/f ratio. For these options, the angle calculation was made in MATLAB® using the Newton-Raphson method. ANN training was carried out in MATLAB® environment with the calculated data. The obtained ANN algorithm is implemented on the STM32F429 microcontroller® and switching signals are obtained. A printed circuit board was designed for the 7-level inverter design and the system was tested on an RL load.

3. System Simulation Analysis

Two different ANN structures were trained for two different harmonic elimination options (3rd - 5th and 5th - 7th). Two different data sets are required for these

trainings. Using the Newton-Raphson method in MATLAB® environment, 151 data were obtained with a modulation index sampling interval of 0.01 for each option in the range of 1.2-2.7 modulation index.

ANN training was carried out using nftool neural network tool in MATLAB® environment. As seen in Figure 4, the general structure of ANNs consists of an input representing the modulation index, eight neurons and three outputs representing the switching angles.

Levenberg-Marquardt algorithm using the least mean square error approach was used to train the network. Mean square error and regression analysis were used as training performance criteria. Figure 5 shows the performance results obtained in both networks, respectively.

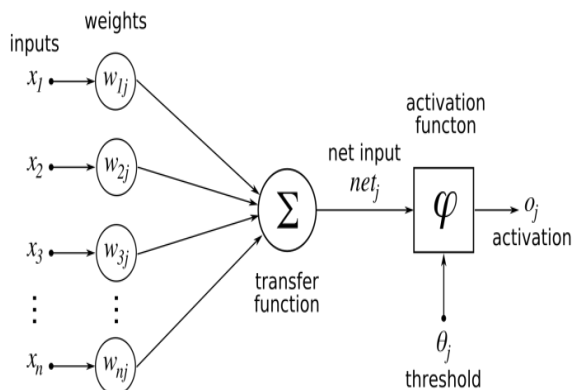


Figure 3. Structure of Single Neuron [13]

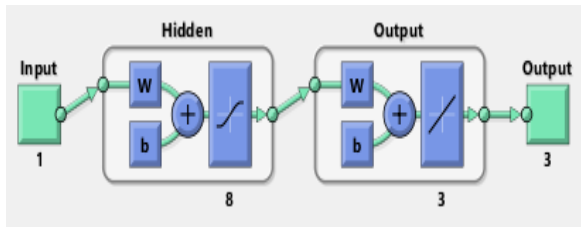


Figure 4. Artificial Neural Network Structure [14]

Results			
	Samples	MSE	R
Training:	105	3.23715e-4	9.99583e-1
Validation:	23	3.47780e-4	9.99546e-1
Testing:	23	3.76651e-4	9.99537e-1

(a)

Results			
	Samples	MSE	R
Training:	121	7.22040e-5	9.99838e-1
Validation:	15	9.51942e-5	9.99772e-1
Testing:	15	2.71884e-5	9.99938e-1

(b)

Figure 5a. ANN Performance Parameters for a) 3rd - 5th Harmonic, b) 5th - 7th Harmonic

The fact that the mean square error is very low and the regression number is very close to 1 means that both networks are well trained. The trained ANNs were compiled in the microcontroller compiler by creating an algorithm to be used on the microcontroller. Figure 6 shows the ANN algorithm.

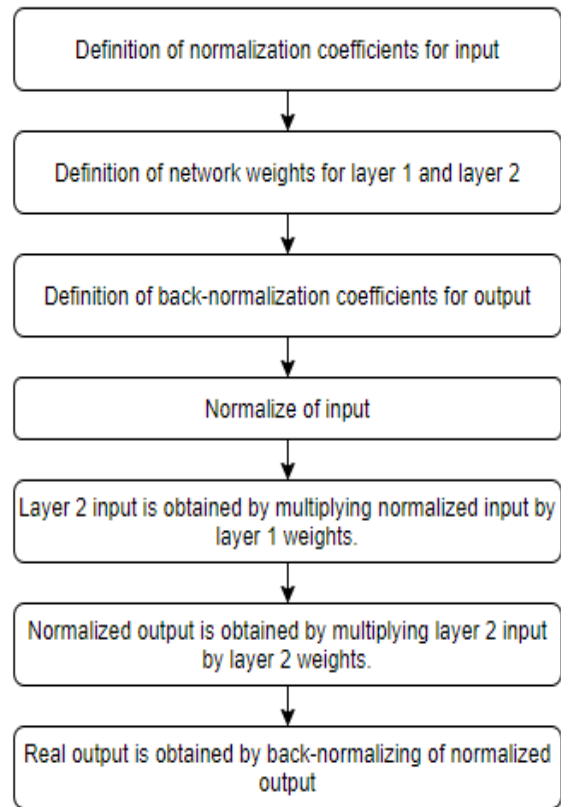


Figure 6. ANN Algorithm

4. Hardware Design

Our inverter is formed by connecting three H-Bridge circuits in series. A total of 4 MOSFETs are used in each circuit, and a gate driver circuit is required to switch these MOSFETs.

However, since the reference voltage for the source leg of the upper MOSFETs is not the ground of the circuit, a technique called bootstrap method is used to increase the voltage between the gate-source to a sufficient level during switching with the used capacitor.

The H-Bridge circuit was designed using the IR2110 driver integrated working with this technique. Figure 7 shows the gate driver circuit for a single module.

This circuit requires three different supply voltages: 25V for the input voltage, 15V and 5V for the IR2110 integrated. For this, other voltages were obtained from 25V input voltage with 7815 (15V) and 7805 (5V) integrations.

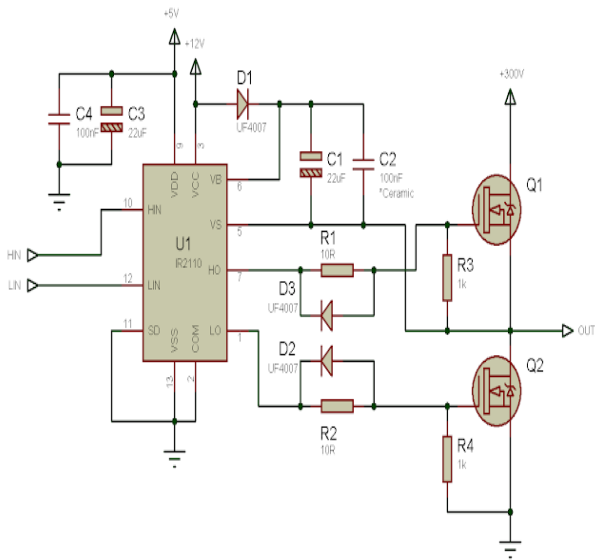


Figure 7. Gate Driver Circuit

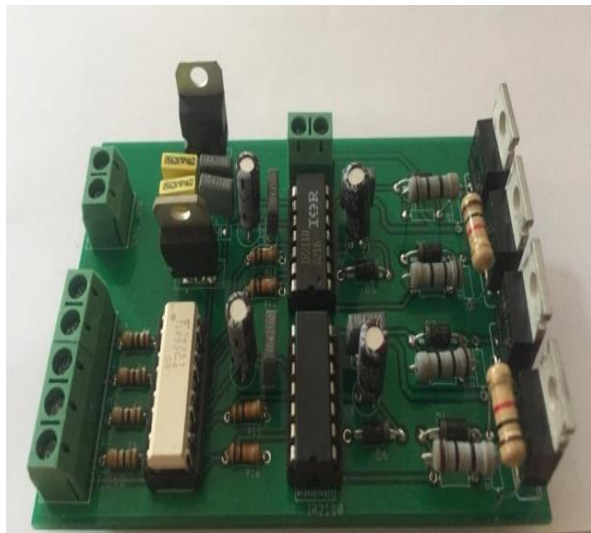


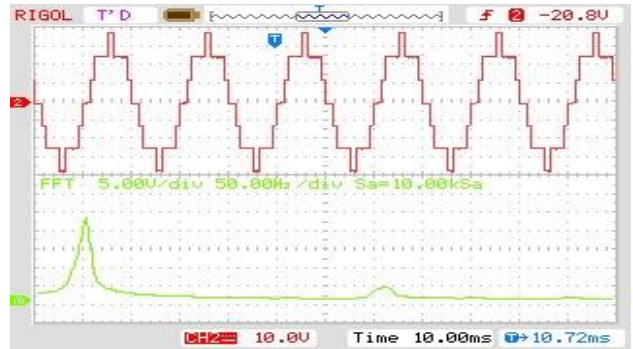
Figure 8. Printed Circuit Board design for H-Bridge

In addition, TLP521-4 integrated circuit is used for isolation between microcontroller and IR2110. Three H-Bridge circuits were obtained by drawing printed circuit in Eagle program.

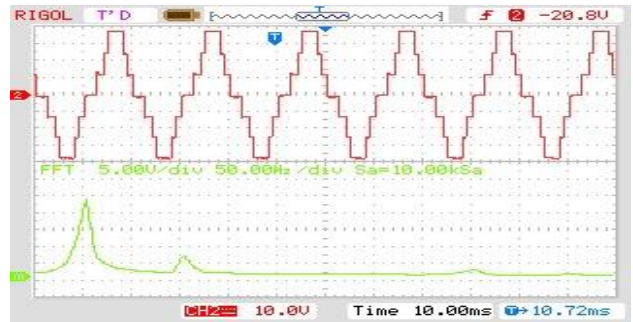
5. Experimental Results

The system was tested with RL load at 20kΩ and 1mH. Three power supplies produce 25V for each circuit, isolated from each other. The output voltages were taken over 10kΩ due to the volt/div restriction of the oscilloscope and half the output voltage (37Vpeak) was measured.

In Figure 9.a, waveforms and fast Fourier transform measurements are given in the modulation index of 2.0 for the situation where the 3rd and 5th harmonics are eliminated, and the 5th and 7th harmonics are eliminated in Figure 9b. As can be seen, the angles provided by both ANNs and the targeted harmonics have been eliminated.



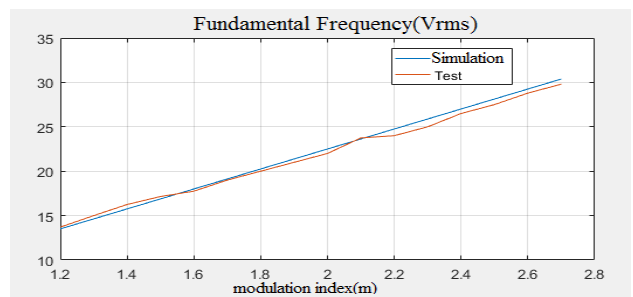
(a)



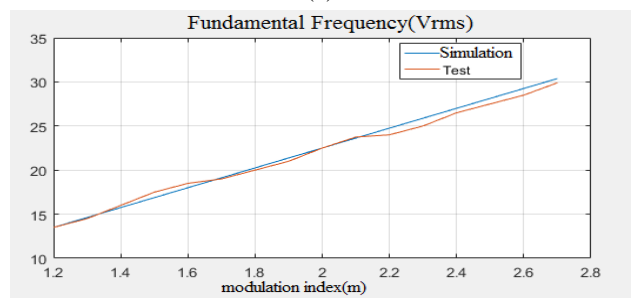
(b)

Figure 9. Wave and FFT Output in a) 3-5 (2.0 Mod. Index) Option, b) 5-7 (2.0 Mod. Index) Option

The fast Fourier transform measurements of the system output were made and the results obtained in the simulation environment were compared with the MATLAB® environment, and the V_{rms} values for the fundamental component and the harmonics they destroy were given as graphs in the range of 1.2-2.7 modulation index. Figure 10 shows the graphics for the fundamental frequency, respectively. As can be seen, with small differences, the result of the experiment follows the simulation result.

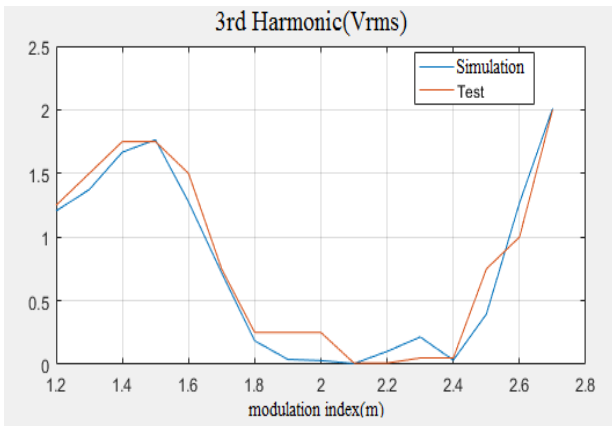


(a)

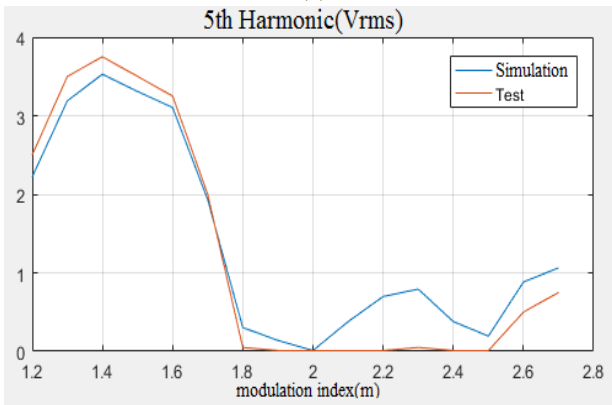


(b)

Figure 10. Comparison of Fundamental Frequency in a) 3-5 Option, b) 5-7 Option

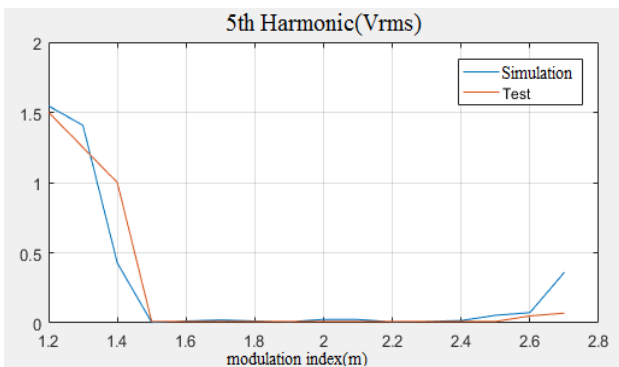


(a)

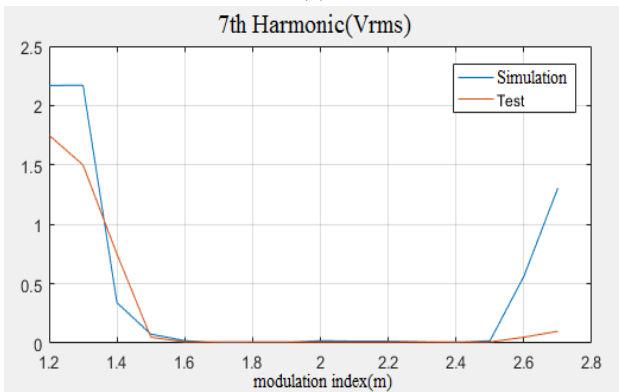


(b)

Figure 11. Comparison of 5th Harmonic in a) 3-5 Option, b) 3-5 Option



(a)



(b)

Figure 12. Comparison of 7th Harmonic in a) 5-7 Option, b) 5-7 Option

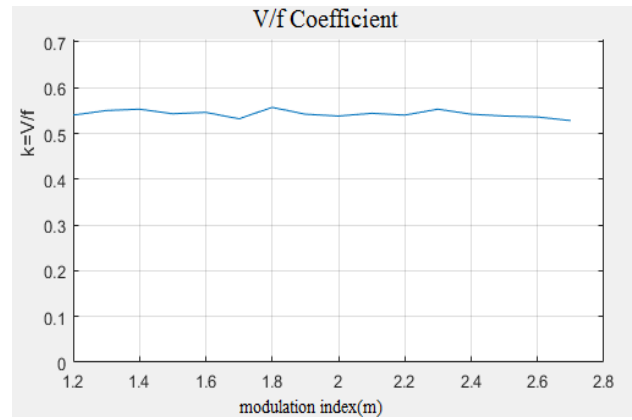


Figure 13. V/f Coefficient in Changeable Frequency Option

Figure 11 shows the graphs obtained for harmonics eliminated in 3-5 option. The experiment results follow the simulation. Full elimination could only be achieved between 1.8-2.5 indices.

In Figure 12, the graphs obtained for harmonics eliminated in 5-7 option are given. Full elimination was achieved between indices 1.5-2.6. In this respect, the situation where 5th-7th harmonics are eliminated provides better performance.

Although there is little difference between the results, in general, the simulation and test results are very close to each other. Since the Newton-Raphson method did not provide results for every modulation index, complete harmonic elimination could not be achieved at certain intervals.

Finally, measurements were taken for V/f control option. As seen in Figure 13, the V/f ratio could be fixed despite very small changes that may occur due to noise and measurement errors.

6. Conclusion

In this study, harmonic elimination has been achieved by using SHE method for 7-level cascade multilevel inverter. The most important contribution of this study was the real-time application of the SHE method using ANN. Offline angle tables used in traditional SHE methods can both exceed the memory and limit the flexibility of the system. A trained ANN has a very fast processing capacity due to its simple structure. Using this feature, angles can be calculated in every period with a speed much higher than the output frequency range of 25Hz-50Hz. In the study, the inverter interface offers two different harmonic options, 3rd – 5th, 5th – 7th and a fixed frequency and variable frequency in V/f ratio.

In addition, the option of changing the output voltage in the 40% -90% modulation range is offered. The weakness of this study is that it cannot achieve complete elimination for each modulation index. To overcome this problem, training data instead of the Newton-Raphson method can be solved by another solution algorithm. As can be seen from the regression value and the results obtained, ANN training is extremely successful.

Considering a new research topic on this project, a system in which the same operations are performed against variable input voltages can be considered.

Declaration

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article. The authors also declared that this article is original, was prepared in accordance with international publication and research ethics, and ethical committee permission or any special permission is not required.

Author Contributions

Ö. Akın developed the methodology and planned the study. İ. Özer worked on artificial neural network design and software implementation. H. Ünlü performed the hardware design and performance analysis.

References

- Santhi, R., Giridharan, K., Kannabhiran, A., *A survey on voltage source multi level inverter topologies*. IEEE International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI), 2017. p. 1262-1271.
- Kar, P. K., Priyadarshi, A., Karanki, S. B., *A modified single phase H-bridge multilevel inverter topology for photovoltaic applications*. in National Power Electronics Conference (NPEC) 2017: Pune, India. p. 340-345.
- Sahu, N., and Londhe, N. D., *Optimization based selective harmonic elimination in multi-level inverters*. in National Power Electronics Conference (NPEC) 2017: Pune, India. p. 325-329.
- Sharma, A., Singh, D., Pandey, V., Gao, S., *Selective harmonic elimination for cascaded H-bridge MLI using GA and NR-Method*. in International Conference on Electrical and Electronics Engineering (ICE3) 2020: Gorakhpur, India. p. 89-94.
- Halim, A. W., Azam, T. N. A. T., Applasamy, K., Jidin, A., *Selective harmonic elimination based on newton-raphson method for cascaded H-bridge multilevel inverter*. International Journal of Power Electronics and Drive Systems (IJPEDS), 2017. **8**(3): p. 1193-1202.
- Chatterjee, A., Rastogi, A., Rastogi, R., Saini, A., and Sahoo, S. K., *Selective harmonic elimination of cascaded H-bridge multilevel inverter using genetic algorithm*. in Innovations in Power and Advanced Computing Technologies (i-PACT) 2017: Vellore, India. p. 1-4.
- Bose, B. K., *Neural Network Applications in Power Electronics and Motor Drives—An Introduction and Perspective*. IEEE Transactions on Industrial Electronics, 2007. **54**(1): p. 14-33.
- Kumar, J., Gambhir, J., Kumar, A., *Control of switching angles for a CMLI using ANN*. in Engineering and Computational Sciences (RAECS) 2014, Panjab University Chandigarh: India. p. 1-6.
- Aravind, P. S., Alexander, S. A., *Harmonic minimization of a solar fed cascaded H bridge inverter using artificial neural network*. in International Conference on Energy Efficient Technologies for Sustainability 2013: Nagercoil, India. p. 163-167.
- Filho, F. J. T., Tolbert, L. M., Ozpineci, B., *Real time selective harmonic minimization for multilevel inverters using genetic algorithm and artificial neural network angle generation*. Proceedings of The 7th International Power Electronics and Motion Control Conference, 2012. p. 895-899.
- Fakhry, M. G., Massoud, A., Ahmet, S., *Quasi seven-level operation of multilevel converters with selective harmonic elimination*, in 2014 26th International Conference on Microelectronics (ICM): Doha, Qatar. p. 216-219.
- Yang, K., Hao, J., and Wang, Y., *Switching angles generation for selective harmonic elimination by using artificial neural networks and quasi-newton algorithm*, in 2016 IEEE Energy Conversion Congress and Exposition (ECCE): Milwaukee, WI. p. 1-5.
- Cybenko, G., *Approximation by superpositions of a sigmoidal function*. Math. Control Signal Systems, 1989. **2**(4): p. 303-314.
- Oluwasegun K. M., Ojo O. A., Ola O. T., Birur A., Cuddy J. and Chan K., *Development of artificial neural network models for predicting weld output parameters in Advanced fusion welding of a magnesium alloy*. American Journal of Modeling and Optimization. 2018. **6**(1): p. 18-34.



Research Article

Effect of aperture averaging on four petal Gaussian beams in atmospheric turbulence

Mert Bayraktar ^{a,*} 

^aTurkish Aerospace Industries, Havacılık Bulvarı No:17, Ankara, Turkey

ARTICLE INFO

Article history:

Received 01 October 2020

Revised 06 February 2021

Accepted 06 March 2021

Keywords:

Aperture averaging
Atmospheric turbulence
Four petal Gaussian beam
Scintillation

ABSTRACT

Aperture averaged scintillation of four petal Gaussian beam is studied in this article. Split step propagation approach which is used in wave propagation applications is selected to model atmospheric turbulence. Results are plotted in two types. First type is the analysis of aperture averaged scintillation versus propagation distance for constant receiver aperture. Second ones involve scintillation performance applying aperture averaging at constant distance. All results are compared with Gauss beam since commercial lasers generally radiates in Gaussian distribution. We observe that four petal Gaussian beam becomes more advantageous under moderate turbulence than weak one. In other point of view, it is possible to obtain less scintillation index by increasing beam order. Our results are applicable optical applications operating in atmosphere.

© 2021, Advanced Researches and Engineering Journal (IAREJ) and the Author(s).

1. Introduction

Because of the recent developments in technology, communication systems with low latency and high data rate need to arise. Related with this, 5G and beyond technologies play a vital role to meet this demand. Free space optics systems are the one of the components of 5G systems. Free space optics systems provide low latency and higher bit rate but performance of these systems is dependent on the atmospheric conditions. One way to overcome the negative effects of atmosphere is beam shaping. Utilizing non-conventional beams, performance of free space optics systems can be improved.

Propagation of different types of beams through random media attracts the attentions of scientists in different points of views. Intensity profiles, scintillation behavior, beam size, and coherency are some of these points [1]. Wave propagation lies in the background of this analysis. Solution of Huygens-Fresnel integral gives the received field through free space or turbulence [2]. For complex source field expressions, split step propagation methods are also used [3]. In some studies, accuracy of this method is also increased [4]. Benefiting from these methods scintillation index which is the most effective noise factor

[5] is measured for untraditional beams. While scintillation index of sine hollow beam is less than Gauss beam [6], less scintillation index than sine hollow beam can be provided by Mathieu-Gauss beam [7]. In addition, we prove that cylindrical-sinc Gaussian beam has less scintillation index than Gauss beam [8]. Besides this this type of beam has a diverging nature propagating in atmospheric turbulence [9]. Considering the effect of scintillation, authors show how to detect information using Gaussian vortex beam [10]. In addition to above beams, there are some other types of beams that provides less scintillation. Regarding with this, it is shown in [11, 12] that Airy and partially coherent Airy beams have less scintillation than Gauss beam. Similarly, scintillation index of truncated flat-topped beam is low [13]. Poynting vector of Weber beam is calculated in [14]. For high frequency wave propagation, the Eulerian Gaussian beam method is generalized in [15].

In addition to methods and some applications listed above, four petal Gaussian beam is introduced by propagating through ABCD system [16]. If Vortex is added to four petal Gauss beam, it is observed that topological charge is quite dominant on the central hollow in far field [17]. While four petal Gaussian vortex beam

* Corresponding author. Tel.: +90-312-811-1800; Fax: +90-312-811-1425

E-mail addresses: mert.bayraktar@tai.com.tr (M. Bayraktar)

ORCID: 0000-0002-0337-7650 (M. Bayraktar)

DOI: 10.35860/iarej.803508

This article is licensed under the CC BY-NC 4.0 International License (<https://creativecommons.org/licenses/by-nc/4.0/>).

shows clockwise phase distribution through the propagation in uniaxial crystal [18], four petal Lorentz-Gauss beam turns into elliptical Gaussian beam in the same medium [19]. M^2 , propagation factor for paraxial region is derived in [20]. Four petal Gaussian beam is propagated through fractional Fourier system and it is concluded as fraction, aperture, and source parameters are effective in received field [21]. Besides these, free space propagation and propagation in turbulent medium is also another interesting field for wave propagation. In the light of this, raise in topological charge provides larger hollow in the center for four petal Lorentz-Gauss vortex beam [22]. Similarly, larger topological charge reduces the effect of atmospheric turbulence on partially coherent elliptic Gaussian vortex beams [23]. Partially coherent four petal Gaussian beam shows resistance to oceanic turbulence if it is generated with larger beam order [24]. Finally, it is shown in [25] that four petal Gaussian vortex beam having incoherency evolves into Gaussian like shape easier if strength of turbulence increases. Airy transform is applied to four petal Gaussian beam and Airy like beam is obtained at the output [26].

Bearing in mind above literature review, we study the aperture averaged scintillation for four petal Gaussian beam under weak and moderate atmospheric turbulent regime. We use split step propagation to model the atmosphere. This model is used for propagate wave through turbulent medium. We analyzed the results with respect to propagation distance and size of receiver aperture. We anticipate that our results will be useful for optical communication and range measurement systems' designers.

2. Source Field Distribution and Method to Measure Performance

2.1 Four petal gaussian beam

Source field expression is written for four petal Gaussian beam [16] as

$$u_s(s_x, s_y) = \left(\frac{s_x s_y}{\alpha^2}\right)^{2n} \exp\left(-\frac{s_x^2 + s_y^2}{\alpha^2}\right) \quad (1)$$

where $n=0,1,2,3,\dots$ refer to beam order, s_x, s_y are transverse plane source coordinates, and $\alpha = \sqrt{\alpha_x^2 + \alpha_y^2}$ being the Gaussian source size. Based on equation.1, Figure 1 and 2 are plotted to show the effect of beam order and Gaussian source size on the initial plane. It is seen from Figure 1 that while beam order increases, hollow region in the center gets larger. When beam order is set to 1, effect of Gaussian source size is seen in Figure 2. For asymmetric case $\alpha_x > \alpha_y$, beam lies along x-axis. Similar idea is valid for $\alpha_y > \alpha_x$. Furthermore, beam spreads on transverse plane symmetrically if α_x, α_y are increased in the same amount.

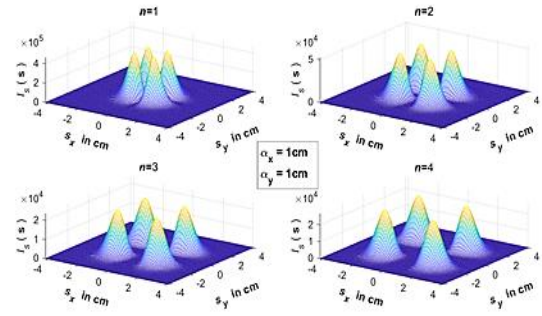


Figure 1. Transverse source plane intensity distribution for constant Gaussian source size

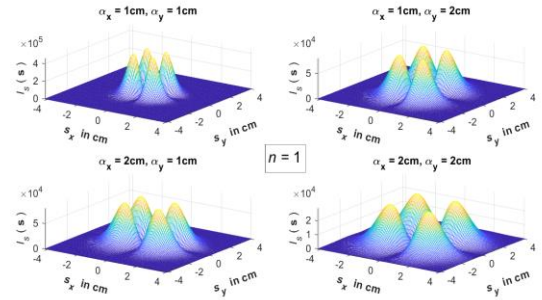


Figure 2. Transverse source plane intensity distribution for constant beam order

2.2. Measurement of scintillation in atmosphere

Numerical split step propagation set-up involves special values presented in table 1 below. In this set-up, between the screens, free space conditions are satisfied. We chose modified von-Karman power spectral density since it is the closest model to Hill spectrum. Refractive index structure constant is the parameter to determine the strength of turbulence and it is taken into account in r_0 . To satisfy the averaging in Eq. 4, number of realizations is taken as 500 which corresponds to infinity in reality. To avoid the complexity of computation of Huygens-Fresnel integral, this method benefits from convolution property of Fourier transform as it is seen from Eq. 2.

After applying all settings, received field after one step is found as

$$u_r(r_x, r_y, L) = A \times F^{-1} \left(F \left[\left(\frac{s_x s_y}{\alpha^2}\right)^{2n} \exp\left(-\frac{s_x^2 + s_y^2}{\alpha^2}\right) \right] \right) F \left\{ \exp \left[\frac{jk}{2L} (s_x^2 + s_y^2) \right] \right\} \quad (2)$$

where

$$A = \frac{-jk \exp(jkL)}{2\pi L} \exp \left[\frac{jk}{2L} (r_x^2 + r_y^2) \right] \exp[\psi(r)] \quad (3)$$

here k denotes wave number, r_x, r_y are receiver plan coordinates, F and F^{-1} refer to Fourier and inverse Fourier transform, and $\psi(r)$ indicates phase fluctuations due to atmosphere and it involves power spectral density. Then, aperture averaged scintillation is evaluated benefiting from received field as:

$$m^2 = \frac{\langle P(r_x, r_y)^2 \rangle}{\langle P(r_x, r_y) \rangle^2} - 1 = \frac{\left(\int u_r(r_x, r_y) u_r^*(r_x, r_y) dS \right)^2 / N}{\left(\int u_r(r_x, r_y) u_r^*(r_x, r_y) dS / N \right)^2} \quad (4)$$

where * indicates complex conjugate. Adapted model of above mathematical equations is presented in Figure 3.

3. Results and Discussions

In this part of the study, we comment on the numerical results.

Table 1. Numerical set-up parameters

Source plane dimensions(S)	10cm X 10cm
Propagation distance(L)	Up to 5500m
Power spectral density	$\phi = 0.023 r_0^{-5} \frac{\exp(-(\frac{f}{f_m})^2)}{(f^2 + f_0^2)^{\frac{11}{6}}}$ [27]
Refractive index structure constant (C_n^2)	For weak turbulence: $10^{-14} m^{-2/3}$ For moderate turbulence : $10^{-13} m^{-2/3}$
Grid size in transverse plane	512 X 512
Inner scale frequency (f_m)	$\frac{5.92}{l_0 2\pi}$, $l_0 \rightarrow 0$
Outer scale frequency (f_0)	$\frac{1}{L_0}$, $L_0 \rightarrow \infty$
Fried parameter	$r_0 = (0.423 (\frac{2\pi}{\lambda})^2 C_n^2 L)^{-\frac{3}{5}}$
Operating wavelength (λ)	1550nm
Number of screens	21: 0-1200m 51: 1200-3200m 91: 3200-5500m
Receiver plane dimensions in Figures 4 and 5	25 X 25 grids
Propagation distance in Figures 6 and 7	3300m

Realization amount for averaging (N)

500

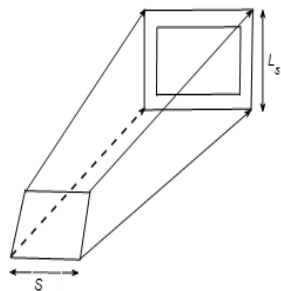


Figure 3. Propagation model through the atmosphere

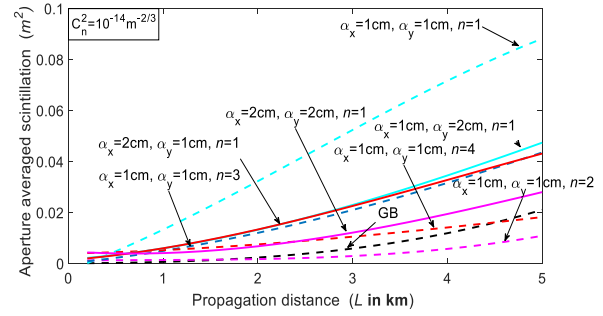


Figure 4. Aperture averaged scintillation of selected beam versus propagation path under weak turbulence

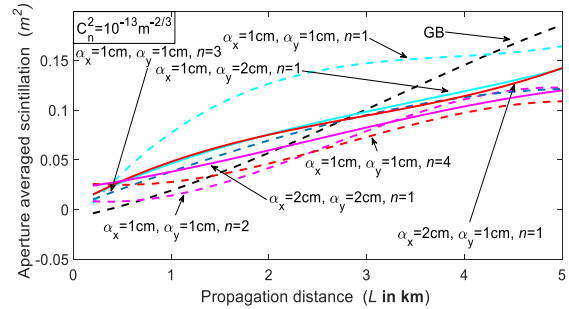


Figure 5. Aperture averaged scintillation of selected beam versus propagation path under moderate turbulence

We show aperture averaged scintillation index variations against propagation distance in Figures 4 and 5. It is seen from these figures that while only second order symmetric beam has advantageous as compared to Gauss beam in weak turbulence, all settings of four petal Gaussian beams have lower scintillation index for stronger turbulence regime.

In moderate turbulence case, scintillation index decreases as beam order raises. Additionally, increase in source size provides similar result with larger beam order. Besides this, asymmetry, without considering the direction, brings scintillation reduction.

In order to show the advance of four petal Gaussian beam, Table 2 is given. As it is observed from this table, however scintillation index of fourth order Four petal Gaussian beam is higher than Gaussian beam at close distance, significant amount of scintillation reduction can be provided by Four petal Gaussian beam at longer distances. At 5 km, scintillation of Gaussian beam is approximately two times higher than Four petal Gaussian beam.

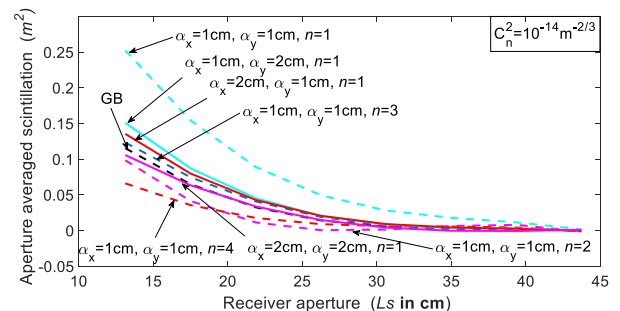


Figure 6. Aperture averaged scintillation of selected beam versus receiver aperture under weak turbulence.

Table 2. Aperture averaged scintillation index under moderate turbulence.

Distance(km)	Gauss beam	Fourth order Four petal Gaussian beam
0.4	0.001086	0.02483
0.6	0.006571	0.02489
0.8	0.01256	0.02586
1	0.01903	0.02767
1.2	0.02592	0.03022
1.4	0.03321	0.03343
1.6	0.04084	0.03723
1.8	0.04879	0.04151
2	0.05071	0.04621
2.6	0.0829	0.06194
3	0.1008	0.07293
3.6	0.1278	0.08853
4	0.1454	0.09734
5	0.1859	0.109

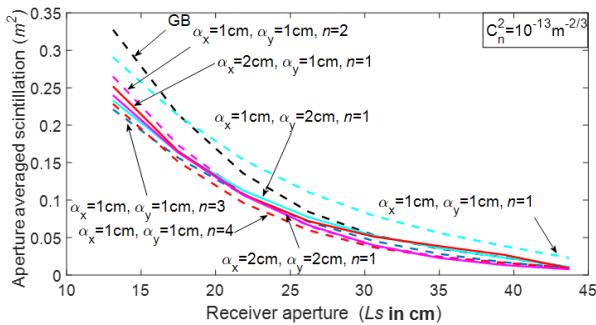


Figure 7. Aperture averaged scintillation of selected beam versus propagation receiver aperture under moderate turbulence

In other point of view, four petal Gaussian beam has lesser scintillation index with versus receiver aperture opening as it is seen from Figures 6 and 7. We see that second and fourth order beam has lesser scintillation index for small receiver apertures. In addition, large source size four petal Gaussian beam has similar performance with Gauss beam under weak turbulence conditions. First order beam has the largest scintillation index. Then asymmetric ones and third order beams follow it.

As it is wanted, scintillation mitigation is seen for moderate turbulence as it is shown in Figure 7. We see that scintillation index of all kinds of four petal beams is less than Gauss beam for small apertures. In advantageous region, from the lowest scintillation index up to the highest one beams are listed as third, fourth order beams, asymmetric and large source size beams, second, and first order beams. We can investigate that four petal beam has less point like scintillation index as compared to Gauss beam since it can be evaluated in small aperture openings. This advantage vanishes when aperture size gets wider.

4. Conclusion

Aperture averaged scintillation of four petal Gaussian beam is studied in this article. Numerical results show that

beam order and scintillation index is inversely proportional to each other. Additionally, scintillation index of four petal Gaussian beam is quite advantageous as compared to Gaussian beam for small aperture openings under moderate turbulence. In addition, four petal Gaussian beam is resistive to atmospheric turbulence. Because, scintillation index of all selected types four petal Gaussian beam is becomes advantageous when turbulence strength raises. Consequently, in the light of these investigations, four petal Gaussian beam can be selected as source beam for optical wireless communication and LIDAR systems operating in turbulence. In the future, our studies will focus on to measure the scintillation of other beams and generation of four petal beam experimentally.

Declaration

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article. The author also declared that this article is original, was prepared in accordance with international publication and research ethics, and ethical committee permission or any special permission is not required.

Author Contributions

M. Bayraktar is responsible for all parts of the study.

Nomenclature

$\alpha = \sqrt{\alpha_x^2 + \alpha_y^2}$: Gaussian source size(m)
C_n^2	: Refractive index structure constant
f_m	: Inner scale frequency
f_o	: Outer scale frequency
F	: Fourier transform
F^{-1}	: Inverse Fourier transform
k	: Wave number (m^{-1})
L	: Propagation distance
L_s	: Receiver plane size
m^2	: Aperture averaged scintillation index
n	: beam order (0,1,2...)
r_x, r_y	: Receiver plane coordinates(m)
s_x, s_y	: Source plane coordinates(m)
S	: Source plane size
u_s	: Source field expression
u_r	: Received field expression
$\psi(r)$: Phase fluctuations
λ	: Operating wavelength

References

1. Wenzel, A.R., *Localization of the Mean and Mean Squared Intensities, and Intensity Fluctuations, of Waves*

- Propagating in a One-Dimensional Random Medium*. Wave Motion, 1985. **7**(6): p. 589-600.
2. Andrews, L.C., *Laser Beam Propagation Through Random Media*. 2 ed. 2005, Washington: SPIE.
 3. Schmidt, J.D., *Numerical Simulation Optical Wave Propagation with examples in MATLAB*. 2010: SPIE.
 4. de Hoop, M.V., J.H. Le Rousseau, and R.S. Wu, *Generalization of the phase-screen approximation for the scattering of acoustic waves*. Wave Motion, 2000. **31**(1): p. 43-70.
 5. Eyyuboglu, H.T. and M. Bayraktar, *SNR bounds of FSO links and its evaluation for selected beams*. Journal of Modern Optics, 2015. **62**(16): p. 1316-1322.
 6. Bayraktar, M., *Estimation of scintillation and bit error rate performance of sine hollow beam via random phase screen*. Optik, 2019. **188**: p. 147-154.
 7. Bayraktar, M., *Scintillation and bit error rate calculation of Mathieu-Gauss beam in turbulence*. Journal of Ambient Intelligence and Humanized Computing, 2020.
 8. Bayraktar, M., *Scintillation and bit error rate analysis of cylindrical-sinc Gaussian beam*. Physica Scripta, 2020. **95**(11).
 9. Eyyuboglu, H.T. and M. Bayraktar, *Propagation properties of cylindrical sinc Gaussian beam*. Journal of Modern Optics, 2016. **63**(17): p. 1706-1712.
 10. Eyyuboglu, H.T., *Optical communication system using Gaussian vortex beams*. Journal of the Optical Society of America a-Optics Image Science and Vision, 2020. **37**(10): p. 1531-1538.
 11. Eyyuboglu, H.T. and E. Sermutlu, *Partially coherent Airy beam and its propagation in turbulent media*. Applied Physics B-Lasers and Optics, 2013. **110**(4): p. 451-457.
 12. Eyyuboglu, H.T., *Scintillation behavior of Airy beam*. Optics and Laser Technology, 2013. **47**: p. 232-236.
 13. Liu, X.L., et al., *Scintillation properties of a truncated flat-topped beam in a weakly turbulent atmosphere*. Optics and Laser Technology, 2013. **45**: p. 587-592.
 14. Rondon-Ojeda, I. and F. Soto-Eguibar, *Properties of the Poynting vector for invariant beams: Negative propagation in Weber beams*. Wave Motion, 2018. **78**: p. 176-184.
 15. Wu, H. and X. Yang, *The Eulerian Gaussian beam method for high frequency wave propagation in the reduced momentum space*. Wave Motion, 2013. **50**(6): p. 1036-1049.
 16. Duan, K.L. and B.D. Lu, *Four-petal Gaussian beams and their propagation*. Optics Communications, 2006. **261**(2): p. 327-331.
 17. Guo, L.N., Z.L. Tang, and W. Wan, *Propagation of a four-petal Gaussian vortex beam through a paraxial ABCD optical system*. Optik, 2014. **125**(19): p. 5542-5545.
 18. Liu, D.J., et al., *Evolution properties of four-petal Gaussian vortex beam propagating in uniaxial crystals orthogonal to the optical axis*. European Physical Journal D, 2015. **69**(9).
 19. Liu, D.J., et al., *Properties of a four-petal Lorentz-Gauss beam propagating in uniaxial crystal orthogonal to the optical axis*. Optik, 2019. **183**: p. 257-265.
 20. Zhou, G.Q. and Y. Fan, *M-2 factor of four-petal Gaussian beam*. Chinese Physics B, 2008. **17**(10): p. 3708-3712.
 21. Tang, B., *Propagation of four-petal Gaussian beams in apertured fractional Fourier transforming systems*. Journal of Modern Optics, 2009. **56**(17): p. 1860-1867.
 22. Deng, W.T., et al., *Four-petal Lorentz-Gauss vortex beam and its propagation in free space*. Optik, 2020. **202**.
 23. Wu, K.N., et al., *Propagation of partially coherent four-petal elliptic Gaussian vortex beams in atmospheric turbulence*. Optics Express, 2018. **26**(23): p. 30061-30075.
 24. Liu, D.J., et al., *Evolution properties of partially coherent four-petal Gaussian beams in oceanic turbulence*. Journal of Modern Optics, 2017. **64**(16): p. 1579-1587.
 25. Liu, D.J., et al., *Propagation properties of partially coherent four-petal Gaussian vortex beams in oceanic turbulence*. Laser Physics, 2017. **27**(1): 016001.
 26. Yaalou, M., Z. Hricha, and A. Belafhal, *Investigation on Airy transform of Four-Petal Gaussian beams*. Optical and Quantum Electronics, 2020. **52**: 165.
 27. Bayraktar, M. and H.T. Eyyuboglu, *Propagation properties of optical bottle beam in turbulence*. Optical Engineering, 2019. **58**(3): 036104.



Research Article

Analysis of whether news on the Internet is real or fake by using deep learning methods and the TF-IDF algorithm

Tilbe Korkmaz^a , Ali Çetinkaya^{b,c,*} , Hakan Aydın^a  and Mehmet Ali Barışkan^a 

^aIstanbul Gelisim University, Faculty of Engineering and Architecture, Computer Engineering, Istanbul, 34310, Turkey

^bIstanbul Gelisim University, Technology Transfer Office Application and Research Center, Istanbul, 34310, Turkey

^cKocaeli University, Faculty of Engineering, Department of Electronics and Communications Engineering, Kocaeli, 41001, Turkey

ARTICLE INFO

Article history:

Received 12 August 2020

Revised 08 October 2020

Accepted 21 October 2020

Keywords:

Natural language processing

Text analysis

Text classification

TF-IDF algorithms

ABSTRACT

Internet use has become increasingly widespread nowadays. In addition, there is a significant increase in the amount of text content produced in digital media. However, the accuracy and inaccuracy of the news we read and the content produced in a large number are also unknown. In this study, classification and analysis of whether the news is real or not were done by using Deep Learning methods. For the English news, the data set created by Katharine Jarmul was used. The data set contained a total of 6336 news items. The distribution of this data set, which consisted of political and political news, was 50% fake and 50% real. The method used in text classification was Term Frequency - Inverse Document Frequency (TF-IDF). The classification was made with the data set used and 93.88% success and 6.12% error were obtained as a result of the analysis.

© 2021, Advanced Researches and Engineering Journal (IAREJ) and the Author(s).

1. Introduction

Especially with the increase of news on the internet, the importance of obtaining accurate news and clear information has also increased. The increase in the number of news has also led to an increase in false of fake news. In this article, text mining methods and deep learning techniques are used together to classify real and fake news. This study is important in terms of automatically and efficiently determining real and fake news, making news analysis faster and more efficient, and eliminating such problems in news analysis.

Deep learning is an artificial intelligence method used in areas such as object recognition, speech recognition, and natural language processing, and it uses multi-layer artificial neural networks. Today, many researchers in different fields such as big data [1-6], autonomous vehicles [7, 8], handwritten character recognition [9, 10], medical image processing, natural language processing [11, 12], signature verification, voice and video recognition, are using deep

learning method in their studies conducted in the most popular and challenging areas of the world [13-16].

In the deep learning method, the learning process is performed through examples. There is no need to use rule sets to solve the problem. It is sufficient to select a model that provides a solution to the problem by evaluating the samples or to create it. Therefore, model selection should be done well [17-19]. Deep learning structures often enable the design, training, and accurate analysis of artificial neural networks using highly programmed interfaces. Some of the most used deep learning libraries are Caffe2, PyTorch, Tensorflow, etc. As a working style, they use multiple graphics processing units (GPU) for high performance; the biggest reason for graphic processors, such as CUDA and cuDNN, to use accelerated libraries is that they can be trained quickly [20].

Today, many studies are carried out on text and sentence classification problems. One of the most important problems of the text classification [21] is that the texts to be classified are not structural [22]. The success rate can be increased by

* Corresponding author. Tel.: +90 212 422 7000 \ 7187; Fax: 0212 422 74 01

E-mail addresses: tilbekorkmaz95@gmail.com (T. Korkmaz), alacetinkaya@gelisim.edu.tr (A. Çetinkaya), haaydin@gelisim.edu.tr (H. Aydın), mabariskan@gelisim.edu.tr (M. A. Barışkan)

ORCID: 0000-0002-6882-5004 (T. Korkmaz), 0000-0003-4535-3953 (A. Çetinkaya), 0000-0002-0122-8512 (H. Aydın),

0000-0002-8039-2686 (M. A. Barışkan)

DOI: 10.35860/iarej.779019

This article is licensed under the CC BY-NC 4.0 International License (<https://creativecommons.org/licenses/by-nc/4.0/>).

using deep learning method in analysis studies conducted with classification algorithms [23] on text processing [24, 25]. As a success criterion, the quality of word representations comes to the fore.

Term Frequency-Inverse Document Frequency (TF-IDF) is used for word representation and Word2Vec [26] is used for fast text [23]. TF-IDF can be used in the so-called subject grouping together with the detection of the articles, in the inference of the author and the classification of the author and the subject [11, 27]. The text classification algorithm developed for author recognition has been successful as a result of experiments [28]. The classification success of texts has been seen as 81.2% [29]. It has been shown that with a 93.3% accuracy rate, much better performance achieved with classification algorithms based on machine learning [9, 10]. In a study on the problem of Turkish text classification [30], as a result of the analysis, it was observed that the stationary words were removed and there was an increase in the classification performance. However, it was not observed that taking word roots had a positive effect on classification accuracy. In other studies, it has been determined that the most successful is the term weights TF and TF * IDF [31].

The purpose is to perform automatic detections by using text mining methods in the interpretations on the internet. In this study, together with 444 comments related to this project, by using the TF-IDF classification algorithm, the success measurements and comparisons ensured in the model can be made. The most successful algorithm among the compared algorithms has been found to be consecutive minimal optimization (88.73%) [32]. It is observed that the number of academic studies published on the internet increases day by day [33]. The words in the news texts are created using the FastText model, which is used very popularly in the recent literature as word embedding. The model is tested individually first by a single sentence and then by the first two sentences through training the full text in each news [34].

A system of detection from comments was developed with analysis on Turkish texts such as social media posts, customer reviews, novels, etc. [35, 36]. The increase in internet and social media use has also increased the sharing of textual information [33]. Classification of the tweets written on the sector has been carried out [37, 38]. It has been investigated that whether there is any theme of emotion among Turkish and English tweets, and if any, a classification of emotions has been performed [39, 40]. The parts obtained from the Turkish texts and the TREMO data set used in the area of emotion extraction have been compared with the classification results obtained using different machine learning algorithms [10]. Classification and modeling are applied to the data set prepared via Facebook. Knowing that it is effective in social media mining, this method has been studied and shown as a text classification [31, 41]. Test data collected using web scraping methods as summaries of news on Turkish news

websites are classified using vector learning methods and depth learning methods together with using a "hot coding". A 90% classification success is achieved [10].

To solve the classification problems of websites by using the artificial neural network [42, 43], two different approaches (dual classification or multiple classifications) are applied. Both approaches were tested on Web sites collected within the scope of a study and a comparison of their performance was observed [44]. ANN is one of the self-learning methods. It leads to the emergence of deep learning [12]. The Regional Based Convolutional Neural Network (RCNN), long and short-term memory (LSTM) [31, 39], and convolutional neural network (CNN) models can be used in sentence classification studies [36].

In the data set used in our deep learning study we conducted classification by using TF-IDF, and 93.88% success and 6.12% errors were detected. Compared to the algorithms used in the references, the success rate was 2.12 percent less than the DSA algorithm and 0.52 percent less than the Random Forest algorithm, but it showed a noticeable success compared to other algorithms.

2. Materials and Methods

2.1 Parts of the System's Control Structure

ANN (artificial neural network) is called logical software working similarly to the human brain and it has been developed to think, learn, generalize, remember, and produce new information like brain. ANN is called synthetic structures that mimic biological neural networks. In the artificial neural network model, there are 3 layers: the input layer, hidden layer, and output layer [45]. The cell structure of the Artificial Neural Network is shown in Figure 1 [46]. The X input, W weights, Σ Transfer function, and net-input that will enter the activation function are shown in figure 1.

As stated in Equation 1, in order to reach the output in Figure 1, the threshold value is summed with the inputs.

$$\text{Result} = f = (\sum_{i=1}^n w_i x_i + \vartheta) \quad (1)$$

Where x_i refers to input values, w_i refers to weight values, and ϑ refers to threshold value.

The activation function provides curved equalization created by the input layers and the output layers.

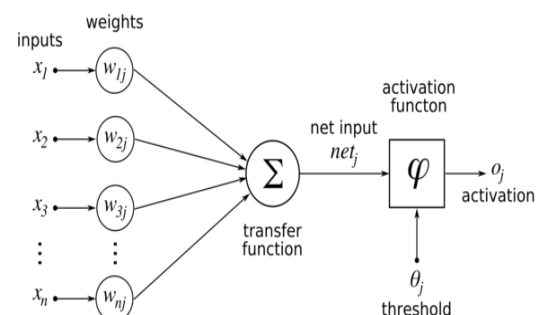


Figure 1. Artificial neural network cell structure

Caution should be exercised when choosing the activation function because this choice significantly affects the performance of ANN. The activation function is unipolar (0-1), but can be selected as bipolar (-1 +1) or linearly qualified. It is the segment that facilitates the introduction and learning of a nonlinear structure in ANN. Sigmoid and ReLU are activation functions.

The most important reason for using the sigmoid activation function is that it compresses the value between 0 and 1. Therefore, it is used in models that find the probability of an event occurring. Although the derivative of the sigmoid function exists, the Sigmoid function is stationary, but its derivative is not stationary. The sigmoid function can cause the model to pause during training. For this reason, it is not recommended for use in machine learning models, which are described as complex. The sigmoid function divides the data into two classes [46]. In Equation 2, the mathematical formula of the Sigmoid function is given. Its values are between 0 and 1.

$$\sigma(X) = \frac{1}{1+e^{-x}} \quad (2)$$

ReLU (Rectified Linear Unit) is a nonlinear function, also known as the transfer function. It is known that the ReLU function takes the value of 0 for the inputs given as negative, whereas it takes the value of x for the inputs given as positive. Today, ReLU is the most used activation function that is generally utilized in Deep Learning and Convolutional Neural Networks [7].

$$R(z) = \max(0, z) \quad (3)$$

The mathematical formula of the ReLU activation function is given in Equation 3. If the value obtained as a result of the activation process is negative, it takes the value 0, if it is positive, it takes the value 1.

Deep learning is expressed as the prediction of data sets created by keeping the created and used data together within the framework of the desired results. It is also called the machine learning method because of the multiple layers created. Deep learning (machine learning) is summarized as the sub-branch of artificial intelligence [13, 47]. The word 'deep' in the Deep Learning method refers to having more than one hidden layer. A sample layer structure of the Deep Learning is shown in Figure 2 [48].

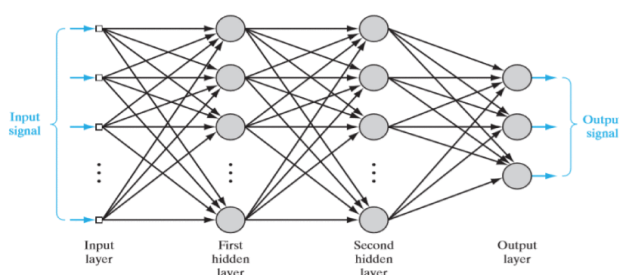


Figure 2. Deep learning sample layer structure

While performing deep learning, three different methods can be used: Supervised, Semi-Supervised, or Unsupervised. In Deep learning methods, together with a large amount of data entering, being able to distinguish between different features are learned. In the learning process, it is determined that the excess data given as input increases success. As the data is processed, it changes its position by passing to the next layer. It is known that as the upper layers are passed, the number of details extracted from the given one increases. It is also known that there are more than one deep model types. These are Multilayer Perceptron, Convolutional Neural Networks, and Recurrent Neural Networks.

Multilayer perceptrons (MLP) have been created as a result of the studies conducted to solve the “exclusive or (XOR)” problem. It has been observed that MLP is effective when classifying and generalizing. Since many inputs will not be enough for a single neuron, more than one neuron is needed for parallel processing. It takes the data from the input layer and transfer to the hidden layer. The intermediate layer can differ as being at least one layer. The exit of each layer becomes the entrance of the next layer. Each neuron is attached to the neurons in the next layer. The output layer determines the output of the network by processing the data received from the previous layer. The number of outputs is equal to the number of elements in the output layer. A training set consisting of sample inputs and outputs is essential for the network to learn.

The usage areas of Deep Learning are increasing day by day. Today, it is mostly used in areas such as face detection, sound detection, and vehicles with auto-pilot feature and called driverless vehicles. In alarm systems, it is thought that since examining the continuous camera recordings leads to loss of time, technologies such as alarm system warnings operating depending on unusual movements provide convenience. Today, this is possible with deep learning methods. Thanks to the deep learning methods used in the health sector, cancer research has gained momentum and it is observed that it eliminates the loss of time. It can be easily diagnosed whether the cells are cancerous or not by recognizing the samples marked as cancerous cells by deep learning algorithms. With this process, fast and successful results are obtained. As in other areas, deep learning methods can be also developed to improve the quality of images in cyber threat analysis.

2.2 Natural Language Processing

In the literature, languages are examined in two categories; these are machine language and the natural language we use in everyday speech. Natural language processing has been developed for computers to understand and interact with the language that people speak. It is also called the process by which computers qualify natural languages.

Directly imported objects cannot be used for the classification of texts. Data is obtained from these extracted

objects and the process is carried out on these data. In the classification process, the separation of the texts into classes by pre-processing the texts ensures that successful results are obtained by facilitating the process to be done in the next step. Various categories are produced on the texts whose classification is desired, and various steps such as Lower-Case process, tokenization process, stop words extraction, stemming process, lemmatization process, and BoW process are applied.

The Lower-Case Process is used to eliminate differences by translating all the words in different sentences into lowercase letters in existing texts.

The Tokenize process stays on the whole of the texts in most data sets. Since it is not functional to operate with these texts, the data must be separated into parts and removed. This process is called token (shred) and Tokenizing is done as a string.

On the other side, words that we use a lot in our daily lives and that does not make sense when used alone are called "Stop Words". The elimination of words such as "and", "or", and "at" in the texts in order not to deviate from the accuracy of the text models to be studied is expressed as the Stop Words process.

The Stemming Process is described as Stemming tightening in Turkish. It is one of the most important steps in data preprocessing. It is known as the process of obtaining a root by trying to cut the suffixes in the front and the end in existing words.

Table 1 shows the examination of the words stemming process. As shown in the table, it does not discriminate while descending to the roots of the words. For this reason, it can be successful in some cases. Lemmatization Process, called as the Lemmatization root analysis in Turkish, is a word unitization. It has similar features with the Stemming process. In both processes, the aim is to obtain root. The Lemmatization Process tries to extract a morphological or semantic root from words. If this process is done, more accurate results will be obtained by discarding the excesses that are described as noise.

Table 2 shows steps of the Lemmatization process. The most important feature of the lemmatization process is that it is linguistic. To obtain an accurate lemma, the given words should be analyzed morphologically.

BoW (Bag of Word) is a process used in natural language processing to simplify models. The words in the texts are kept in a bag with the BoW model. While holding together, grammar and sequence errors are not taken into account. It is the most used method for extracting attributes in texts in data sets being used.

Table 1. Stemming process

Word	Suffixes	Stemming
Fly	-es	Fli
Flying	-ing	fly

Table 2. Lemmatization process

Word	Morphological Information	Lemmatization
Flies	3rd person, Singular, Present Simple Tense	fly
Flying	3rd person, noun-verb, Present Tense	fly

We know that Natural Language Processing is examined in two categories; machine and everyday speech languages. Natural language processing is used because computers cannot understand the language that people speak. It is known that the texts written on computers are digitized as 0 and 1. The large data sets used in Natural Language Processing and the texts in these data sets are digitized by various methods by computers. This digitization process is carried out with various algorithms and methods.

Loss function is defined as function that measures the error and success rate of the created model. The last layer of artificial neural networks is known as the layer where the loss function is defined. The Loss function transforms the error rate calculation problems that occur during training into an optimization problem. The Loss function first calculates the difference between the actual values of the estimation made by the training model. Since the prediction will not be good, the difference between the real values and the estimated values will be large if the necessary care is not taken in the model creation. Accordingly, the loss value will be high. In cases where the model is created well, the difference will decrease, and in cases where it is the same, the loss will be 0.

Optimization (Optimizer) Algorithms are methods used to minimize the difference between the output value and the actual value produced at the output layer in the artificial neural network. Based on the size of the data set used in model training, there are 3 different types of algorithms described as slope descent.

Metrics are measurements where data can be counted and numerical values are obtained. The values obtained as a result of these measurements are expressed as total or ratio. Metrics are used when showing all or different elements of a dimension.

Term Weighting is a process performed on the roots obtained by natural language processing. If the term exists in the document, a weighted value of this term is formed. However, if this value is not available, it is expressed as 0.

In terms of the Term Frequency (TF), the weighting process is calculated in the document where the data is found. Weighting is made according to the term in the document. Since it takes value as much as amount of it in the document, the more it passes, the more value it gets.

Inverse Document Frequency (IDF) weighting gives weight according to the number of documents in which a term is mentioned. In the total documents used, the low presence of the same term increases the distinguishing

feature, namely the IDF value. Otherwise, if a term is detected more than once in many documents, there is a decrease in the IDF value as its distinctiveness decreases. The result to be reached here is a standardization obtained by multiplying the frequency of the terms used with the IDF value.

$$w_{i,d} = tf_{i,d} \times \log(n/df_i) \quad (4)$$

Equation 4 represents the calculation for the term i and the document d . TF is calculated first. The ratio of the number of times the term is mentioned in the document to the most mentioned term is shown in equation 5 below.

$$tf_{i,d} = \frac{f_{i,d}}{\max f_{i,d}} \quad (5)$$

$$\{0,1\} \quad (6)$$

Binary looks for the presence of terms in used documents. It is shown in Equation 6.

$$f_{t,d} \quad (7)$$

Raw Frequency refers to the number of repetitions of terms in documents and the total number of words in the documents. It is shown in Equation 7.

$$\log(1 + f_{t,d}) \quad (8)$$

Log Normalization is taking its logarithm based on the value of $f_{t,d}$ calculated in Equation 7. It is shown in Equation 8.

$$0.5 + 0.5 \frac{f_{t,d}}{\max f_{t,d}} \quad (9)$$

As a result of weighting, Double Normalization generates a value between 0.5 and 1. It is shown in Equation 9. By dividing the value obtained by calculating Raw Frequency to the maximum number of terms in the document, the ratio of the other terms is calculated and standardization is performed in frequency.

$$K + (1 - K) \frac{f_{t,d}}{\max f_{t,d}} \quad (10)$$

Double Normalization K is calculated as a smoothing term that can be viewed as scaling with the largest TF value. It is

shown in Equation 10.

The Confusion Matrix is used to measure performance in classification algorithms performed on test data and where the accuracy of the actual values is known.

2.3. Flow Chart of The System

In this study, using Deep Learning methods, the classification and analysis of the news, which were real or fake, were done. The flow diagram of the system is shown in Figure 3. After the program is run, it connects to the data set we obtain as ready. The data set consists of real/fake news in English. After this news is classified as real and fake, the pre-processing of text begins. In the context of this study, various text preprocessing were carried out. Space vectors created with TF-IDF, the digitization process of texts was started. Before starting the model training, the training and test data in the data set were parsed. After the model training starts, real values and predictive values analysis were performed. As a result of this analysis, the realization steps of the application were completed by drawing the graph of the values formed.

3. Applications

Almost everyone is exposed to fake news, which is still present today. This type of news can be described as information pollution and they are misleading. In this study, to combat news that does not reflect reality, a model that could make a distinction between real and fake news by using deep learning methods and the text classification process was created. Analyses were made between the trained data and the test data, and these analyses were visualized through graphics.

In the study, Python and Python Libraries were used as the programming language because it is open source. The data (Katharine Jarmul) [49, 50] includes 6336 different news (political, political). The distribution of the data set is 50% fake and 50% real. The fact that the data set is distributed evenly affects the performance criterion positively.

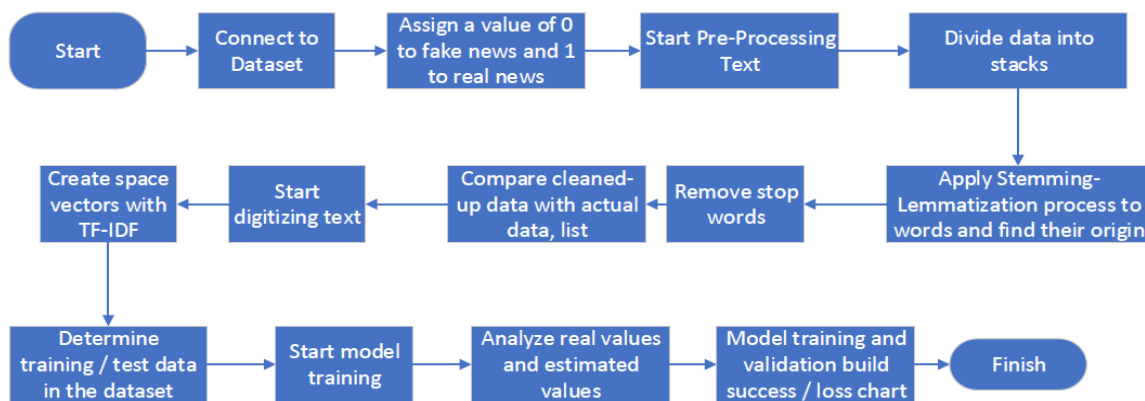


Figure 3. Flow chart of the system

As can be seen in Figure 4 and 5 respectively, real and fake news were determined in the training data, and whereas the value of 1 was assigned to real news, the value of 0 was assigned to fake news.

After completing the definition of training data in the data set, the texts were cleaned so that the texts could be classified correctly. The first step in text preprocessing is the fragmentation of data that stacks together. In Figure 6, the raw version of the news headlines is shown without preprocessing.

The visually cleared data are shown in Figure 7. To get more successful results in model training, the parsed texts were involved in the cleaning process. These processes were carried out to convert all letters to lowercase letters, to remove special characters (!, ^.%), to remove numbers, to divide texts into parts (tokenize), to delete Stop Words (ineffective words), and finally to find the morphological root of the words.

After the text preprocessing was completed, the Vector Space Model, which was one of the methods used in digitizing the texts, was used to bring the texts into digital form. Although each word represents a vector, when they are put together, the MxN-sized term-document matrix is formed. In this matrix, M is the number of news and N is the number of terms. The reason for the use of TF-IDF (Term Frequency - Inverse Document Frequency) management is to calculate the term weights, the frequency of the terms used with TF (Term Frequency), and the amount of all news in the data set checked with IDF (Inverse Document Frequency).

The text digitization process was completed after this stage. Before starting modeling, it is necessary to make the necessary definitions to build the network. To classify news texts, 3 layers were created in the model. Although the first and second layers were hidden layers, the number of nodes in each layer was 32. The third layer was called the output layer. The number of nodes in this layer was 1. ReLU (Rectified Linear Unit) was used as the activation function in the layers called the first and second hidden layers.

In model training, a layer called Dropout was added to prevent the model from memorizing the data. This layer sets the specified input data to zero.

```

title ... label
4490 State Department says it can't find emails fro... REAL
8062 The 'P' in PBS Should Stand for 'Plutocratic' ... FAKE
8622 Anti-Trump Protesters Are Tools of the Oligarc... FAKE
4021 In Ethiopia, Obama seeks progress on peace, se... REAL
4330 Jeb Bush Is Suddenly Attacking Trump. Here's W... REAL
[5 rows x 3 columns]
    
```

Figure 4. Identifying real and fake news

```

title ... type
6330 State Department says it can't find emails fro... 1
6331 The 'P' in PBS Should Stand for 'Plutocratic' ... 0
6332 Anti-Trump Protesters Are Tools of the Oligarc... 0
6333 In Ethiopia, Obama seeks progress on peace, se... 1
6334 Jeb Bush Is Suddenly Attacking Trump. Here's W... 1
[5 rows x 4 columns]
    
```

Figure 5. Value assignment to real and fake news

The purpose of this synchronization is to prevent the model from overfitting the data. The dropout value was chosen between 0 and 1. In the third layer, the Sigmoid function was used as an activation function. These functions were used to describe nonlinear relationships in the model and for binary classification problems. If these functions were not used, there would be a decrease in learning achievement in the model.

The distinction between fake / real news used in the model is called the binary classification problem. Many loss functions are used in binary classification problems. In this study, binary_crossentropy was used as the loss function. To minimize the value between real and estimated values, the rms probe algorithm was used as an optimizer algorithm, and accuracy values were monitored and measured.

While model training is performed, it is not right to put all the data into training for the training to be more successful. It needs to be broken down into specific pieces and involved in training. For this reason, data are divided into pieces called chunks. The number of times that all data will pass through Artificial Neural Networks is determined by Epoch (cycle). In this study, the data obtained from the model was defined in 20 Epoch and the model training was completed with 128 stacks for the data to be passed. The Scikit-Learn library was used to visualize the analysis between the values created as a result of the training and the actual values.

Index	title	text	label	type
0	You Can Smell Hillary's Fear	And it's an interesting question. Hillary's old strategy was to lie and deny that the FBI even had ...	FAKE	0
1	Watch The Exact Moment Paul Ryan Committed Political Suicide At A Trump Rally (VIDEO)	- AOL NEWS POLITICS (@AOLPOLITICS) November 3, 2016 The Democratic Party couldn't have asked for a better moment of f... The ringing endorsement of the man he clearly hates on a personal...	FAKE	0
2	Kerry to go to Paris in gesture of sympathy	Among roughly 40 leaders who did attend was Israeli Prime Ministe...	REAL	1
3	Bernie supporters on Twitter erupt in anger against the DNC: 'We tried to warn you!'	- Ana Navarro (@ananavarro) November 9, 2016 Popular left-wing Facebook page The Other 98%, which was pro-Sand...	FAKE	0
4	The Battle of New York: Why This Primary Matters	Trump needs to capture more than 50 percent of the vote statewide...	REAL	1
5	Tehran, USA	being tormented because of my ethnicity and religion in those ear... Through patience, humor, and understanding, I was able to offer m... In 1998, I became special assistant to the Master Chief Petty Off...	FAKE	0

Figure 6. Display of the headlines before text preprocessing

Indi	Type	Size	Value
0	str	1	smell hillari fear
1	str	1	watch exact moment paul ryan commit polit suicid trump ralli video
2	str	1	kerri pari gestur sympathy
3	str	1	berni support twitter erupt anger dnc tri warn
4	str	1	battl york primari matter
5	str	1	tehran usa

Figure 7. Data cleared as a result of text preprocessing

```

Epoch 1/20
3819/3819 [=====] - 1s 368us/step - loss: 0.6218 -
accuracy: 0.8023 - val_loss: 0.5344 - val_accuracy: 0.8894
Epoch 2/20
3819/3819 [=====] - 1s 288us/step - loss: 0.4307 -
accuracy: 0.9194 - val_loss: 0.3732 - val_accuracy: 0.9059
Epoch 3/20
3819/3819 [=====] - 1s 292us/step - loss: 0.2803 -
accuracy: 0.9440 - val_loss: 0.2696 - val_accuracy: 0.9082
Epoch 4/20
3819/3819 [=====] - 1s 288us/step - loss: 0.1847 -
accuracy: 0.9612 - val_loss: 0.2121 - val_accuracy: 0.9176
Epoch 5/20
3819/3819 [=====] - 1s 291us/step - loss: 0.1217 -
accuracy: 0.9736 - val_loss: 0.1830 - val_accuracy: 0.9224

```

Figure 8. The result of the study

4. Results and Discussion

The result of the study is shown in Figure 8. The news in the dataset used was determined to be fake and real. Before the training, the value of 0 (zero) was assigned to fake news and the value of 1 (one) was assigned to real news. The above figure shows the first 5 stages of 20 epoch training. It is seen that the loss values are decreasing here. It is observed that as the training increases, the loss value decreases. On the other hand, the increase in value accuracy rates as a result of the training can be seen, and the software can distinguish the news as fake or real. As the training increases, the val_accuracy value increases, that is, the proximity to 1 increases, and the distinguishing of true news from false news can be ensured.

Deep learning contributes to the development of artificial intelligence through the use of machine learning with many applications. The technology, which develops day by day,

allows new technological developments to occur to prevent the increase of false information in the digital environment. Natural language processing is a field developed to transform the information, which is increasing day by day in the digital environment, into a language that can be understood by machines. The increase in data day by day makes the processing of these data difficult. Natural language processing categorizes the texts and classifies them according to their characteristics, which allows us to quickly respond to the desired data.

In this study, using deep learning methods, news described as real or fake were classified by the TF-IDF (Term Frequency - Inverse Document Frequency) method. Keras library was used for model training and Pandas library was used to process data. Training and test data were determined just before starting the model training. A 3-layer model was created for model training. The number of training for the model was determined as 20.

As a result of the study, as shown in Figure 9, while the training loss of the model decreases in every step of the training, the success of the trained model increases. This increase is provided by the success of the model in training. Accordingly, when it comes to the 20th Epoch, while the decrease in the training loss of the model approaches to zero, the inversely proportional training loss increases.

In Figure 10, when we set the number of steps for our model training as 200, it is seen that the training loss decreases to zero, and accordingly, the loss of validation increases as the model is trained.

As a result of the statistical information we obtained in model training, the training of the model was calculated as 92.00% and the margin of error was 8.00%.

As seen in Figure 11, the number of epoch used for our model training was chosen as 10000. As a result of the statistical information we obtained in model training, the training of the model was calculated as 93.88% and the margin of error was 6.12%. A Confusion matrix was obtained by using the Scikit-Learn library to evaluate the model used.

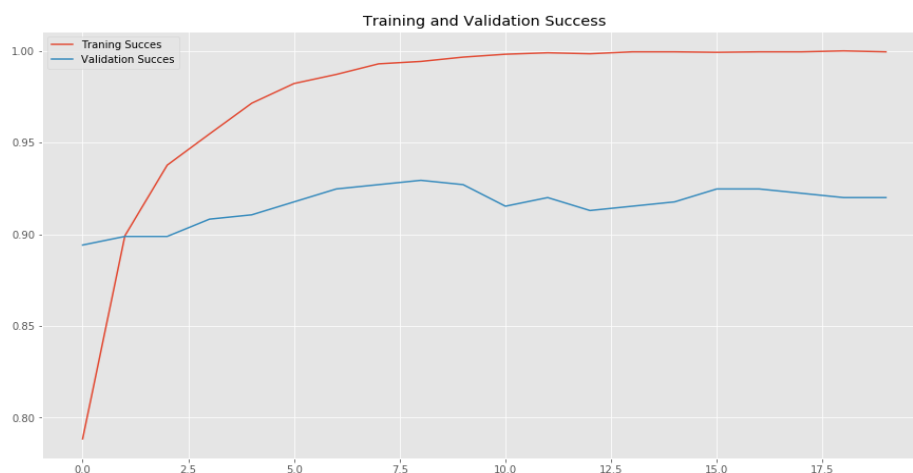


Figure 9. Model training success and loss in 20 Epoch

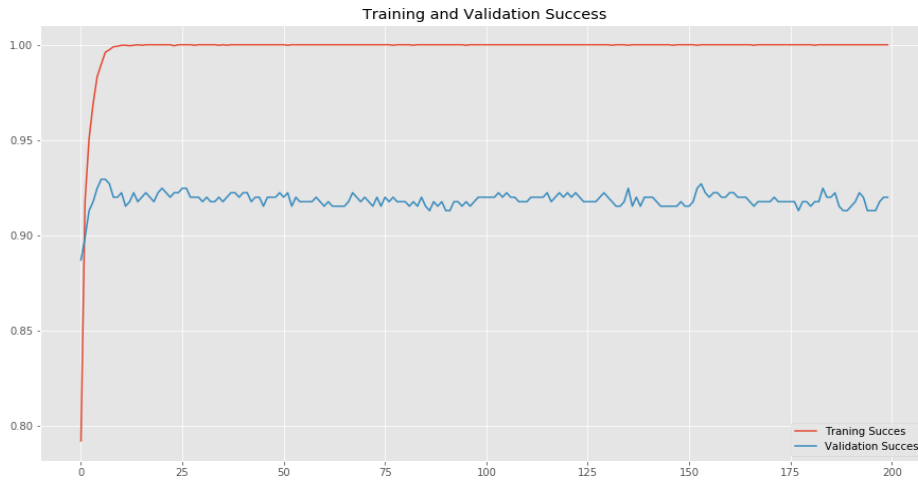


Figure 10. Model training success and loss in 200 Epoc

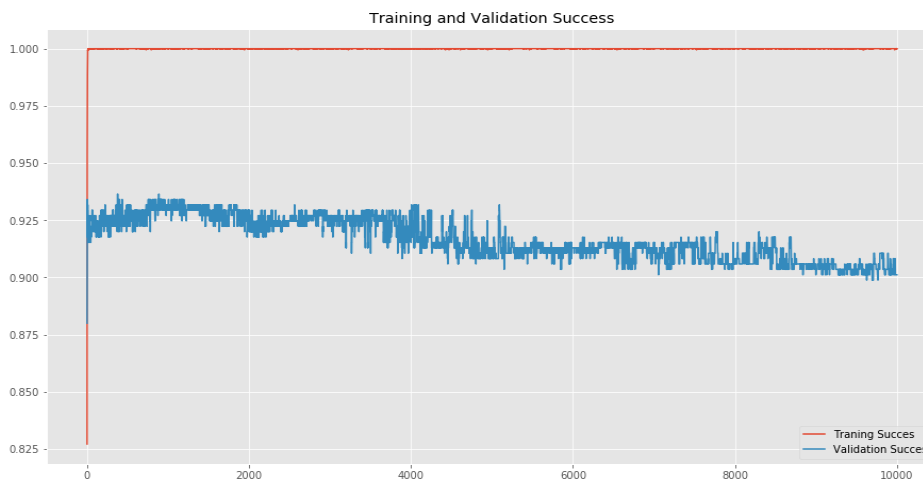


Figure 11. Model training success and loss in 10000 Epoc

The Confusion Matrix is given in Figure 12. It facilitated the observation of the result of the real values and performed on test data. It is seen that 79 of the 929 fake news selected as test data are real, and 85 of the 998 real news are determined as fake. Table 3 contains the success and error rates of the experiments conducted during the software development process. 8 experiments were carried out during the study process.

Experiments were carried out on the same data sets. Training and test data were not changed and necessary studies were carried out on the same model. The difference between experiments is that the epoch values during training were different. It is seen that as the Epoch value increases, the rate of success from training increases. While the success rate of 20 epochs in Experiment 1 was 91.06%, the error rate was 8.94 %. In Experiment 2, while the success rate of 200 epoch was 92.00 %, the error rate was 8.00 %. In Experiment 3, 500 epochs result in 92.89 % success rate and 7.11 % error rate. It was observed that as the number of epochs increased, the success rate increased and the error rate decreased.

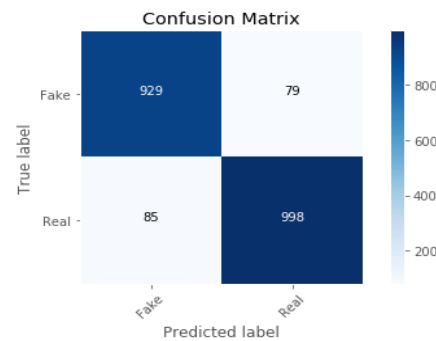


Figure 12. Confusion matrix

Table 3. Experiments and their results

Experiment No.	Epoch Value	Success Rate	Error Rate	Times (Sec)
1	20	91.06%	8.94%	24
2	200	92.00%	8.00%	241
3	500	92.89%	7.11%	595
4	1000	93.00%	7%	1216
5	2000	93.15%	6.85%	2532
6	3000	93.46%	6.54%	3795
7	5000	93.65%	6.35%	6290
8	10.000	93.88%	6.12%	12730

Table 4. Evaluation of the studies in the literature

Study No.	Technique Used	Algorithm Used	Success Rate	Fail Rate
Bilgin et al., 2019	Deep Learning	DSA algorithm	96.00	4.00
Bilgin, 2019	Deep Learning	Random Forest Algorithm	94.40	5.60
This Study	Deep Learning	Tf-Idf Algorithm	93.88	6.12
Süzen, 2019	Deep Learning	Convolutional Neural Networks (CNN)	93.70	6.30
Yücel et al., 2018	Machine Learning	Sequential Minimal Optimization Algorithm	88.00	12.00
Arı et al., 2017	Deep Learning	Convolutional Neural Networks (CNN)	87.30	12.70
Özmen et al., 2019	Deep Learning	it is not specified.	81.20	18.80

The data used in Table 4 were obtained as a result of the literature review and they were brought together to compare the algorithm and methods used. Based on these data, it was concluded that the study with the highest success was realized with the deep learning technique and the DSA algorithm [4]. It is observed that the lowest success percentage was obtained in the study performed with deep learning techniques [28]. Compared to other algorithms, the success rate of the method applied within the scope of the study is over 90%. The study has a low success rate of 2.12 percent according to the DSA algorithm. There is a 0.52 percent success rate reduction compared to the random Forrest algorithm. A noticeable increase in success has been achieved compared to the other algorithms in Table 4.

The abundance of information available today increases day by day with the advancement of technology. It is also easier for people to access information thanks to the internet. More research is carried out to ensure the accuracy of the information obtained. Therefore, the importance of deep learning, which is one of the sub-branches of artificial intelligence, is increasing day by day. Natural language processing used in deep learning techniques enables the processing of existing texts. Accordingly, classification of texts in different categories according to their subjects and processes becomes easier. We carried out our study by taking into account the access to the real news and the speed of reaching it accordingly.

Before starting our studies, various literature studies were reviewed and 6 different studies were selected by referring to 50 different studies. By making various comparisons over the data sets used, the common features of these selected studies were analyzed taking into account the differences in the methods used and the success achieved.

Based on the study we carried out, the data set, working model, and methods used in word analysis will change depending on the technological developments and more successful results can be obtained.

The fact that the information that needs to be confirmed is increasing day by day leads to an increase in the number of these types of studies. By adding a Turkish data set to the study we have carried out, an interface can be created and a website that is constantly updated and renews itself can be

created using the necessary web services.

5. Conclusions

In this study, whether the news found on the internet is real or fake was analyzed using deep learning methods and the TF-IDF algorithm. In this way, by identifying and detecting real news with high success rates, it can be ensured that people can distinguish between real and fake news; in addition, by filtering fake news before news is added to search engines, only real news can be shown to the user. Considering this, a system was developed that can easily and quickly reveal the accuracy of news without having to be read by the user. With this system, real news can be shown to the user at a rate of 93.88%. Using the developed method, social media analysis, comment analysis, and twitter analysis can be carried out on web pages and accuracy analysis can be performed in different areas. In addition, by analyzing comments made on applications used in mobile platforms, this method can also be improved in a way to detect fake comments.

Declaration

The authors declare no potential conflicts of interest with respect to the research, authorship, and/or publication of this article. The authors also declare that this article is original, has been prepared in accordance with international publication and research ethics, and ethical committee permission or any special permission is not required.

Author Contributions

T. Korkmaz, A. Çetinkaya and H. Aydın have planned and designed the research. M. A. Barışkan contributed to the process of working stages. T. Korkmaz, A. Çetinkaya and H. Aydın have carried out data collection and analysis. All the authors discussed the results and contributed to the final form of the article.

References

1. Arroyo-Fernández, I., Méndez-Cruz, C. F., Sierra, G., Torres-Moreno, J. M., and Sidorov, G., *Unsupervised sentence representations as word information series:*

- Revisiting TF-IDF*. Computer Speech & Language, 2019. **56**: p. 107-129.
2. Kim, D., Seo, D., Cho, S., and Kang, P., *Multi-co-training for document classification using various document representations: TF-IDF, LDA, and Doc2Vec*. Information Sciences, 2019. **477**: p. 15-29.
 3. Koyun, A., and Afşin, E., *Derin öğrenme ile iki boyutlu optik karakter tanıma*. Türkiye Bilişim Vakfı Bilgisayar Bilimleri ve Mühendisliği Dergisi, 2017. **10**(1): p. 11-14.
 4. Yapıcı, M. M., Tekerek, A., and Topaloğlu, N., *Literature review of deep learning research areas*. Gazi Mühendislik Bilimleri Dergisi (GMBD), 2019. **5**(3): p. 188-215.
 5. Patil, T., Pandey, S., and Visrani, K., *A review on basic deep learning technologies and applications*. In Data Science and Intelligent Applications, 2020. pp. 565-573. Springer.
 6. Calisan, M., and Sakar, C. O., *Classification of short-texts by utilizing an external knowledge source*. Journal of Science and Engineering, 2017. **19**(57).
 7. Grigorescu, S., Trasnea, B., Cocias, T., and Macesanu, G., *A survey of deep learning techniques for autonomous driving*. Journal of Field Robotics, 2020. **37**(3), 362-386.
 8. Küçük, D., and Arıcı, N., *Doğal dil işleme derin öğrenme uygulamaları üzerine bir literatür çalışması*. Uluslararası Yönetim Bilişim Sistemleri ve Bilgisayar Bilimleri Dergisi, 2018. **2**(2): p. 76-86.
 9. Acı, Ç., and Çırak, A., *Türkçe haber metinlerinin konvolüsyonel sinir ağları ve word2vec kullanılarak sınıflandırılması*. Bilişim Teknolojileri Dergisi, 2019. **12**(3): p. 219-228.
 10. Ertam, F., *Sosyal medya verileri için etkili bir sınıflandırma yaklaşımı*. Fırat Üniversitesi Mühendislik Bilimleri Dergisi, 2017. **29**(2): p. 67-73.
 11. Sjarif, N. N. A., Azmi, N. F. M., Chuprat, S., Sarkan, H. M., Yahya, Y., and Sam, S. M., *SMS spam message detection using term frequency-inverse document frequency and random forest algorithm*. Procedia Computer Science, 2019. **161**, p. 509-515.
 12. MATLAB Deep Learning. [cited 2020 20 May]; Available from: <https://www.mathworks.com/solutions/deep-learning.html>
 13. Şeker, A., Diri, B., and Balık, H. H., *Derin öğrenme yöntemleri ve uygulamaları hakkında bir inceleme*. Gazi Mühendislik Bilimleri Dergisi, 2017. **3**(3): p. 47-64.
 14. Hark, C., Uçkan, T., Seyyarer, E., and Karacı, A., *Metin özetlemesi için dijital merkezliklerine dayalı denetimsiz bir yaklaşım*. Bitlis Eren Üniversitesi Fen Bilimleri Dergisi. 2019. **8**(3): p. 1109-1118.
 15. Güldal, H., and Çakıcı, Y., *Ders yönetim sistemi yazılımı kullanıcı etkileşimlerinin sınıflandırma algoritmaları ile analizi*. Journal Of Graduate School Of Social Sciences, 2017. **21**(4).
 16. Salur, M. U., Aydın, İ., and Karaköse, M. *gömülü derin öğrenme ile tehdit içeren nesnelere gerçek zamanda tespiti*. Dümf Mühendislik Dergisi, 2019. **10**(2): p. 497-509.
 17. Machine learning platform. [cited 2020 7 May]; Available from: <https://www.tensorflow.org>
 18. Bayram, F., *Derin öğrenme tabanlı otomatik plaka tanıma*. Politeknik Dergisi, 2020. **23**(4): p. 955-960,
 19. Sakarya, Ş., and Yılmaz, Ü. *Derin öğrenme mimarisini kullanarak b130 indeksinin tahmini*. European Journal Of Educational And Social Sciences, 2019. **4**(2): p. 106-121.
 20. PyTorch Deep learning library. [cited 2020 20 May]; Available from: <https://pytorch.org>
 21. Yıldırım, S., and Yıldız, T., *Türkçe için karşılaştırmalı metin sınıflandırma analizi*. Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi, 2017. **24**(5): p. 879-886.
 22. Akdoğan, Ö., and Özel, S. A., *Effects of feature extraction techniques on classification of turkish texts*. Çukurova Üniversitesi Mühendislik-Mimarlık Fakültesi Dergisi, 2019. **34**(3): p. 95-108.
 23. Tezgider, M., Yıldız, B., and Aydın, G., *Improving word representation by tuning word2vec parameters with deep learning model*. In 2018 International Conference on Artificial Intelligence and Data Processing (IDAP), 2018. pp. 1-7. IEEE.
 24. Çınar, A., *Veri madenciliğinde sınıflandırma algoritmalarının performans değerlendirilmesi ve r dili ile bir uygulama*. Öneri Dergisi, 2019. **14**(51): p. 90-111.
 25. Özmen, E. P., and Özcan, T., *Dolandiricilik tespiti üzerine melez sınıflandırma ve regresyon ağacı uygulaması*. Yönetim Bilişim Sistemleri Dergisi, 2019. **5**(2): p. 12-20.
 26. Dündar, E. B., and Alpaydın, E., *Learning word representations with deep neural networks for turkish*. In 2019 27th Signal Processing and Communications Applications Conference (SIU), 2019. pp. 1-4. IEEE.
 27. Yücel, A., and Köylü, M. K., *Spam içerikli e-postaların tespiti için bir metin madenciliği uygulaması: terimlerin gama ilişkisi katsayısına dayalı polarizasyonu*. Uluslararası Yönetim Bilişim Sistemleri Ve Bilgisayar Bilimleri Dergisi, 2018. **2**(2): p. 95-104.
 28. Bilgin, M., *Türkçe metinlerin sınıflandırma başarısını artırmak için yeni bir yöntem önerisi*. Uludağ Üniversitesi Mühendislik Fakültesi Dergisi, 2019. **24**(1): p. 125-136.
 29. Demir, Ö., Baban Chawai, A., and Doğan, B., *Türkçe metinlerde sözlük tabanlı yaklaşımla duygu analizi ve görselleştirme*. PORTA. 2019; **1**(2): p. 58-66.
 30. Ertam, F., *Deep learning based text classification with Web Scraping methods*. In 2018 International Conference on Artificial Intelligence and Data Processing (IDAP), 2018. pp. 1-4. IEEE.
 31. Thakkar, A., and Chaudhari, K., *Predicting stock trend using an integrated term frequency-inverse document frequency-based feature weight matrix with neural networks*. Applied Soft Computing, 2020. 106684.
 32. Göker, H., and Tekedere, H., *Fatih projesine yönelik görüşlerin metin madenciliği yöntemleri ile otomatik değerlendirilmesi*. Bilişim Teknolojileri Dergisi, 2017. **10**(3): p. 291-299.
 33. Safali, Y., Nergiz, G., Avaroğlu, E., and Doğan, E., *Deep learning based classification using academic studies in doc2vec model*. In 2019 International Artificial Intelligence and Data Processing Symposium (IDAP), 2019. pp. 1-5. IEEE.
 34. Karakoç, E., and Yılmaz, B., *Deep learning based abstractive turkish news summarization*. In 2019 27th Signal Processing and Communications Applications Conference (SIU), 2019. pp. 1-4. IEEE.
 35. Doğan, O., *Türkiye’de veri madenciliği konusunda yapılan lisansüstü tezler üzerine bir araştırma*. Gazi Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi, 2017. **19**(3): p. 929-951
 36. Pirana, G., Sertbaş, A., and Ensari, T., *Sentence classification with deep learning method for virtual assistant applications*. In 2019 3rd International Symposium On Multidisciplinary Studies And Innovative Technologies (Ismsit), 2019. pp. 1-5. IEEE.
 37. Ayata, D., Saraçlar, M., and Özgür, A., *Turkish tweet sentiment analysis with word embedding and machine*

- learning*. In 2017 25th Signal Processing and Communications Applications Conference (SIU), 2017. pp. 1-4. IEEE.
38. Nasim, Z., and Haider, S., *Cluster analysis of urdu tweets*. Journal of King Saud University-Computer and Information Sciences, 2020.
 39. Bilgin, M., and Şentürk, İ. F., *Danışmanlı ve yarı danışmanlı öğrenme kullanarak doküman vektörleri tabanlı tweetlerin duygu analizi*. Balıkesir Üniversitesi Fen Bilimleri Enstitüsü Dergisi, 2019. **21**(2): p. 822-839.
 40. Alpkoçak, A., Tocoglu, M. A., Çelikten, A., and Aygün, İ., *Türkçe metinlerde duygu analizi için farklı makine öğrenmesi yöntemlerinin karşılaştırılması*. Dokuz Eylül Üniversitesi Mühendislik Fakültesi Fen Ve Mühendislik Dergisi, 2019. **21**(63): p. 719-725.
 41. Süzen, A. A., *LSTM derin sinir ağları ile üniversite giriş sınavındaki matematik soru sayılarının konulara göre tahmini*. Engineering Sciences, 2019. **14**(3): p. 112-118.
 42. Altan, G., *DeepGraphNet: grafiklerin sınıflandırılmasında derin öğrenme modelleri*. Avrupa Bilim ve Teknoloji Dergisi, 2019. (Özel Sayı): p. 319-327.
 43. Doğan, F., and Türkoğlu, İ., *Derin öğrenme modelleri ve uygulama alanlarına ilişkin bir derleme*. Dicle Üniversitesi Mühendislik Fakültesi Mühendislik Dergisi, 2019. **10**(2): p. 409-445.
 44. Chouseinoglou, O., and Şahin, İ., *Metin madenciliği, makine ve derin öğrenme algoritmaları ile web sayfalarının sınıflandırılması*. Yönetim Bilişim Sistemleri Dergisi, 2019. **5**(2): p. 29-43.
 45. Partal, T., Kahya, E., and Ciğizoglu, K., *Yağış verilerinin yapay sinir ağları ve dalgacık dönüşümü yöntemleri ile tahmini*. İtüdergisi/D, 2008. **7**(3): p. 73-85.
 46. Arı, A., and Berberler, M. E., *Yapay sinir ağları ile tahmin ve sınıflandırma problemlerinin çözümü için arayüz tasarımı*. Acta Infologica, 2017. **1**(2): p. 55-73.
 47. Özcan, B., Kumru, P. Y., and Fırlı, A., *Forecasting operation times by using artificial intelligence*. International Advanced Researches and Engineering Journal, 2018. **2**(2): p. 109-116.
 48. Al Qasem, O., Akour, M., and Alenezi, M. *The influence of deep learning algorithms factors in software fault prediction*. IEEE Access, 2020. **8**, 63945-63960.
 49. Fake News Dataset. [cited 2020 15 May]; Available from: <https://www.datacamp.com/community/tutorials/scikit-learn-fake-news>
 50. Khan, J. Y., Khondaker, M., Islam, T., Iqbal, A., and Afroz, S., *A benchmark study on machine learning methods for fake news detection*. 2019. arXiv preprint arXiv:1905.04749.



Research Article

Recombinant production of *Thermus aquaticus* single-strand binding protein for usage as PCR enhancer

Özlem Kaplan ^{a,*} , Rizvan İmamoğlu ^b , İskender Şahingöz ^c  and İsa Gökçe ^c 

^aIstanbul University, Department of Molecular Biology and Genetics, Istanbul, Turkey

^bBartın University, Department of Biotechnology, Bartın, 74110, Turkey

^cTokat Gaziosmanpaşa University, Department of Genetic and Bioengineering, Tokat, 60250, Turkey

ARTICLE INFO

Article history:

Received 13 July 2020

Revised 21 September 2020

Accepted 01 October 2020

Keywords:

Polymerase Chain Reaction,
Single Strand Binding Protein,
Thermus aquaticus

ABSTRACT

Single-stranded DNA-binding (SSB) proteins play an important role in DNA metabolism involving DNA replication, recombination, and repair in all living beings. In molecular biology, SSB proteins are used as enhancers to increase the efficiency and specificity of PCR. Thermostable SSB protein eliminates secondary structure or dimer formation and significantly increase the effectiveness of amplification of DNA fragments. In this study, it was ensured that the SSB gene of thermophilic bacteria *Thermus aquaticus* (*T. aquaticus*) was cloned into the pET28b vector and expressed in *E. coli* BL21 (DE3) PLYSE cells. Then, the purification of the SSB protein produced in *E. coli* BL21 (DE3) PLYSE cells was performed. 20 mg SSB protein was obtained from 1L bacterial culture, and its purity was more than 90%. It was shown by the PCR experiment that the SSB protein produced in this study could increase the amplification efficiency.

© 2021, Advanced Researches and Engineering Journal (IAREJ) and the Author(s).

1. Introduction

PCR (polymerase chain reaction) technique is one of the most basic techniques of molecular biology. PCR protocols are used in a wide range such as routine diagnosis, genomic, and transcriptomic analysis, and these protocols need to be developed [1]. Due to the intrinsic properties of template DNA used in PCR such as high GC content and tendency to form secondary structures, PCR products do not occur under standard reaction conditions. This is a factor that limits the routine use of PCR. Strategies that can provide low-cost and reliable reaction conditions are needed for large scale PCR experiments [2]. In general, template DNAs contain long homopolymer regions, high GC content, and tandem repeats; therefore, it is difficult to amplify the template DNA by PCR. DNA templates with more than 65% GC content give very weak signals when observed under standard PCR conditions, and non-specific product formation is observed [3]. PCR can be improved by

making some changes in reaction conditions. For example, “Touch-Down PCR”, performed by decreasing the annealing temperature step by step in each cycle, and “Hot Start PCR” using modified DNA polymerases cause serious improvement in PCR results [2]. In addition, when various enhancers such as tetramethylammonium chloride (TMAC), dimethyl sulfoxide (DMSO), Betaine, Glycerol, Formamide, non-ionic detergents, and their combinations are added in PCR, these enhancers increase efficiency, specificity, and reproducibility of the PCR amplification. It is particularly effective in ensuring the specificity of formamide and DMSO PCR products. In particular, formamide and DMSO are effective in ensuring the specificity of PCR products. Betaine can reduce the T_m value of DNA and it is effective in DNA amplification with a long and high GC content [4]. It also increases PCR's product efficiency and detection sensitivity. Often, 2 or more PCR enhancers are used together to make a PCR reaction work better. In addition

* Corresponding author. Tel.: +0212-440-00-00 (15111)

E-mail addresses: ozlem.kaplan@istanbul.edu.tr (Ö. Kaplan), rizvanimamoglu@gmail.com (R. İmamoğlu), iskender675@gmail.com (İ. Şahingöz), isa.gokce@gop.edu.tr (İ. Gökçe)

ORCID: 0000-0002-3052-4556 (Ö. Kaplan), 0000-0002-6306-4760 (R. İmamoğlu), 0000-0003-4127-4772 (İ. Şahingöz), 0000-0002-5023-9947 (İ. Gökçe)

DOI: 10.35860/iarej.766741

This article is licensed under the CC BY-NC 4.0 International License (<https://creativecommons.org/licenses/by-nc/4.0/>).

to the known classic effect enhancers, new chemicals and substances are also being investigated. Trehalose, homoectoine, Zn^{2+} -1,7-bis(4-quinolylmethyl)-1,4,7,10-tetraazacyclododecane (Zn^{2+} -Q2-cyclen), and some nanoparticles are among the investigated chemicals. For example, trehalose, which displays a function like betaine, can facilitate PCR of GC-rich DNA by reducing the T_m value and DNA polymerase. Thanks to trehalose, the long PCR process can be also improved [5]. Homoectoine, which is a derivative of 1-ectoine, is more effective than betaine. Homoectoine reduces the T_m value and can increase the specificity of PCR at concentrations lower than betaine [6]. Zn^{2+} -Q2-cyclen, which can be specifically bound to deoxythymidine (dT), disrupts the hydrogen bond between adenine and thymine, decreases T_m value, and is effective in increasing PCR specificity [7]. However, despite their potential to greatly increase PCR effectiveness, commercial enhancers have significant disadvantages, such as the cost and unknown composition [2].

DNA-binding proteins of bacteriophage T4, such as gp32 T4 and SSB protein from *Escherichia coli* (EcoSSB), significantly increase the effectiveness of amplification of DNA fragments [8]. SSBs bind to single-stranded DNA and protect it from the digestion of nuclease. It ensures that ssDNA remains in a suitable conformation in DNA replication, repair, and recombination processes. In addition, SSB protein can physically interact with some proteins that play a role in this DNA metabolism. Therefore, it can be said that SSBs also play an important role in DNA metabolism [9, 10]. Since thermostable SSB proteins bind without denaturing the primers, they eliminate secondary structure or dimer formation. Thermostable SSB proteins are also highly effective in increasing the effectiveness of PCR when PCR conditions are considered [11]. In particular, SSB protein prevents primer dimers in multiplex PCR studies carried out with primers that have different annealing temperatures. In the studies conducted so far, the SSB protein of many thermophilic bacteria has been produced and their roles in increasing the effectiveness of PCR have been revealed [12-20].

The thermostable SSB of all bacteria belong to the *Deinococcus-Thermus* phylum except for SSB from *Thermoanaerobacter tengcongensis* [17]. They have been found in *T. thermophilus* [12], *T. aquaticus* [12], *D. geothermalis* [13], *D. murrayi* [14], *D. radiopugnans* [15], *D. radiodurans* [18], *D. grandis*, and *D. proteolyticus* [19].

Dabrowski et al. showed that the SSB protein of *T. aquaticus* is highly effective in providing amplification of weakly amplified regions by conducting experiments with a wide variety of DNA templates [12]. The SSB protein of *T. aquaticus*, which is a thermophilic bacteria,

contains 266 amino acids and its molecular weight is 30 kDa [21, 22].

E. coli, which is frequently used in recombinant protein production, has advantages such as low cost and rapid production of recombinant proteins. Many proteins are produced by the recombinant DNA technology using *E. coli* strains [23].

In this study, *T. aquaticus* SSB gene used as a PCR enhancer was cloned into the pET28b vector. Expression of *T. aquaticus* SSB (*TaqSSB*) protein was performed in *E. coli* BL21 (DE3) *PlysE* cells. Thereafter, *TaqSSB* protein was produced and purified. It was shown that the purified *TaqSSB* protein could be used as a PCR enhancer.

2. Material and Method

2.1. Cloning of *TaqSSB* gene into the pET28b vector

The SSB gene sequence of *T. aquaticus* was amplified by PCR using the primers SBBBamHISense 5'TTTTGGATCCAATGGCTCGAGGCCTGAAC3', SSBHindIIIReverse 5'TTTTAAAGCTTTCAAACGGCAAATCCTCCTC 3'. Primers are designed using the *TaqSSB* gene nucleotide sequence (AF276705) in NCBI (National Center for Biotechnology Information). Sense primer has BamHI restriction cutting site and reverse primer has HindIII cutting site.

PCR was performed using 50 ng template DNA, 10 mM dNTP mix, 10mM sense primer and reverse primer, 10 X Pfu Polymerase PCR buffer, and 1 U Pfu DNA polymerase that had a final volume of 50 μ l. The program of the PCR device was set as follows: first, 2 min 1 cycle at 95°C; then, a total of 31 cycles, including 1 min at 95°C for denaturation, 1 min at 55°C for annealing, and 1 min at 72°C for extension; and 5 min at 72°C for the final extension. The obtained PCR products were purified with the PCR products cleaning kit and digested with BamHI and HindIII restriction enzymes. The pET28b plasmid to be used for cloning was also cut with the same restriction enzymes. *TaqSSB* gene and pET28b plasmid digestion by restriction enzymes were ligated with T4 DNA ligase enzyme at room temperature for 16 hours after purification performed with PCR products cleaning kit. The ligation products were transferred to *E. coli* DH5 α cells and spread on LB medium containing kanamycin (50 mg/ml). Plasmid DNA isolation was done from the colonies, and diagnostic restriction digest and diagnostic PCR were performed on plasmids. The obtained products were analyzed in 1% agarose gel (Figure 1 and Figure 2, respectively).

2.2. Production and Purification of *TaqSSB* protein

For the expression of N terminal hexahistidine-tagged (6xHis) *TaqSSB* protein, plasmid DNAs from positive

clones (pET28bSSB) were transferred to *E. coli* BL21 (DE3) *PLysE* cells. Transformed cells were inoculated into 50 ml LB medium containing kanamycin (50mg/ml) and chloramphenicol (34 mg/ml) and induced by IPTG when OD_{600} : 0.6. Before and after induction performed with IPTG, the total cellular protein was analyzed in 12% SDS-PAGE (Figure 3).

E. coli BL21 (DE3) *PLysE* cells producing *Taq*SSB protein were incubated for 3 hours at 240 rpm at 37°C after the induction with IPTG. Then, the cells were collected by centrifugation at 8000 rpm for 5 minutes. The cells were dissolved in lysis buffer (100 mM sodium phosphate, 100 mM NaCl, and pH 7.8); then, PMSF (100 mM) and Benzamidine (100 mM) were added and lysed on ice by a sonicator. The cell lysate was kept at 95 °C for 20 minutes and then centrifuged at high speed for 60 minutes at 30 000 rpm. Purification of the SSB protein in the supernatant was carried out with the Ni-NTA column thanks to His-tag added to the protein. 100 mM sodium phosphate, 100 mM NaCl, and pH 7.8 were used as purification buffers. Elution of the protein from the column was carried out using 100 mM sodium phosphate, 100 mM NaCl, 300 mM imidazole, and pH 7.8 buffer [24]. The obtained protein was analyzed in 12% SDS-PAGE (Figure 4) and its amount was determined by the Bradford method.

2.3. Usage of *Taq*SSB protein for PCR amplification

Different concentrations of the purified SSB protein (50 ng/μl, 100 ng/μl, 250 ng/μl,) were added to the PCR mixture. Plasmid DNA containing the proteinase K (proK) gene was used as a template in PCR. The PCR was carried out using proK specific primers (ProteinazKNDEI sense TTTTCATATGGCTGCGCAGACCAACGCTCCTT and ProteinazKHINDIII reverse TTTTAAAGCTTTCAAGCCTGGTAGTTGTTGTA).

The program of the PCR device was set as follows: first, 2 min 1 cycle at 94°C; then, a total of 34 cycles, including 30 seconds at 94°C for denaturation, 1 min at 60°C for annealing, and 30 seconds at 72°C for extension; and 5 min at 72°C for final extension. PCR products were analyzed in 1% agarose gel electrophoresis (Figure 5).

3. Results and Discussion

3.1. Cloning of *Taq*SSB gene into the pET28b vector

*Taq*SSB gene was cloned into the pET28b vector. Confirmation of cloning was first performed by validation restriction cut using BamHI and HindIII restriction enzymes. As seen in Figure 1, it is revealed with the presence of the SSB gene region around 800 bp and the pET28b vector around 5300 bp.

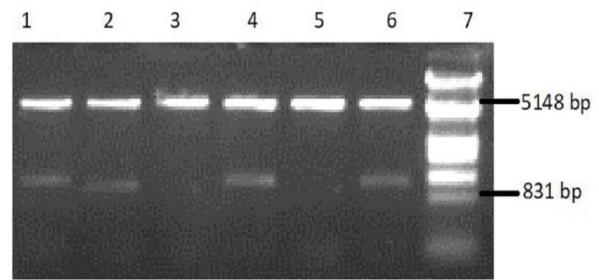


Figure 1. Analysis of the validation restriction cut result of the SSB gene cloned into the Pet28b vector in 1% agarose gel. 1, 2, 4, 6 are pET28b vector containing the SSB gene (positive clones pET28bSSB plasmids). 3, 5 are pET28b vector without the SSB gene and 7 is λDNA/EcoRI/HindIII Marker

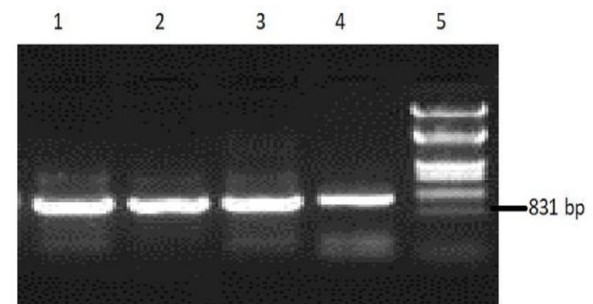


Figure 2. Analysis of the validation PCR result of the SSB gene cloned into the Pet28b vector in 1% agarose gel. 1, 2, 3, 4 are PCR products using pET28b plasmids (pET28bSSB) containing the SSB gene as a template in Figure 1. 5 is λDNA/EcoRI/HindIII Marker

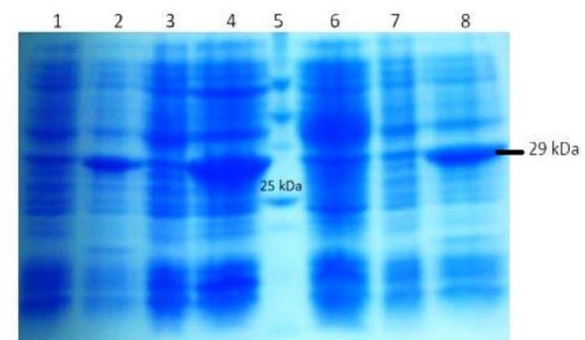


Figure 3. Analysis of the expression of *Taq*SSB protein in *E. coli* BL21 (DE3) *PLysE* cells in 12% SDS-PAGE. 1, 3, 7 are *E. coli* BL21 (DE3) *PLysE* cells containing the pET28bSSB plasmid before induction with IPTG. 2, 4, 8 are *E. coli* BL21 (DE3) *PLysE* cells containing the pET28bSSB plasmid after the IPTG induction. 6 shows *E. coli* BL21 (DE3) *PLysE* cells without pET28bSSB plasmid after the induction with IPTG. 5 is BioRad dual color precision plus protein marker

Diagnostic PCR was also performed by using plasmids that were found to be positive as a result of the diagnostic restriction digestion. As expected, the band belonging to the SSB gene, which appeared around 800 bp, demonstrated the verification of cloning (Figure 2).

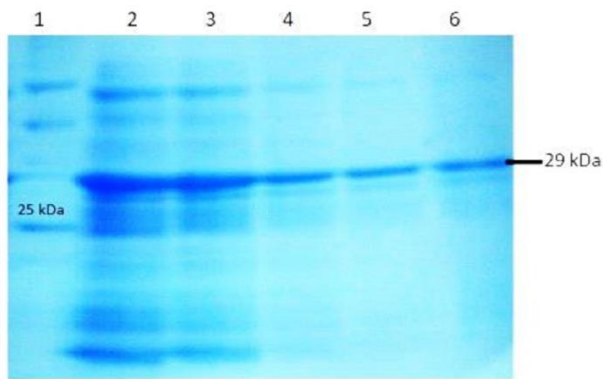


Figure 4. Analysis of the purified *TaqSSB* protein in 12% SDS-PAGE. 1: BioRad dual color precision plus protein marker, 2-6: Fractions of the purified *TaqSSB* protein

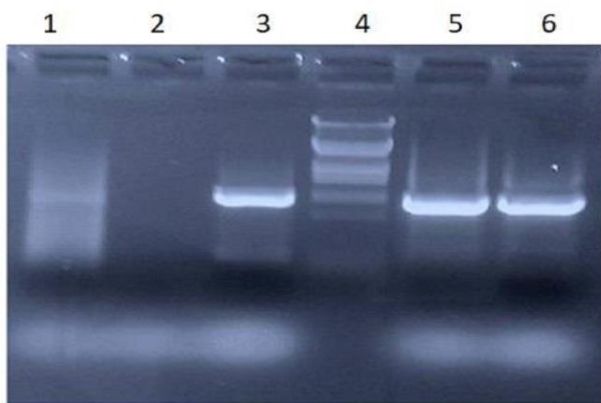


Figure 5. Analysis of the *TaqSSB* protein usage efficiency in PCR in 1% agarose gel. 1: Amplification of the *proK* gene without SSB protein. 3, 5, and 6: Amplification of the *proK* gene performed by adding various amounts of SSB protein; it contains SSB protein at concentrations of 50 ng/μl, 100 ng/μl, 250 ng/μl, respectively. 4: λDNA/EcoRI/HindIII Marker

3.2. Production and Purification of *TaqSSB* protein

The pET28b plasmid containing the SSB gene, which was confirmed to be cloned due to diagnostic restriction digest and diagnostic PCR, was transferred to *E. coli BL21 (DE3) PlySE* cells. For the analysis of protein expression, the total cell lysate before and after induction with IPTG was displayed in 12% SDS-PAGE. As seen in Figure 3, the expression of approximately 29 kDa size SSB protein was observed at the location where it was expected after induction with IPTG.

6xHis-*TaqSSB* protein was produced in *E. coli BL21 (DE3) PlySE* cells and purified by nickel affinity chromatography. The fractions of the purified *TaqSSB* protein analyzed in %12 SDS-PAGE (Figure 4), and its amount was determined by the Bradford method. As a result, 20 mg SSB protein was obtained from 1L bacterial culture and its purity was over 90%.

3.3. Usage of *TaqSSB* protein for PCR amplification

The *TaqSSB* protein in the 6th well shown in Figure 4 was taken in various volumes and added to the PCR reaction. PCR products were analyzed on 1% agarose gel. In Figure 5, it is shown that the *TaqSSB* protein, produced and purified recombinantly in this study, is very effective in increasing the effectiveness of PCR.

4. Conclusion

PCR is a powerful molecular biology technique. Various enhancers and their combinations are used to solve technical problems occurring in PCR. These enhancers increase the specificity, efficiency, and overall effectiveness of PCR. Traditionally used PCR enhancers (betaine, TMAC, formamide, DMSO) help solve complex secondary structure formation in GC-rich DNA templates. These enhancers can reduce the melting temperature of the primers as well as DNA templates [2]. SSB, an important protein for in vivo DNA replication, is effective in shortening PCR extension time and increasing PCR detection sensitivity [25]. SSB prevents primer dimer formation and increases PCR specificity [11]. Considering the temperature conditions in PCR, it can be said that SSBs isolated from thermophilic bacteria are very effective in PCR. Various studies have demonstrated the role of SSB proteins of thermophilic bacteria in increasing PCR activity [12-17]. *TaqSSB* protein provides highly effective amplification on a wide variety of weakly amplified DNA templates [12]. The use of SSB protein in the PCR technique have been routinely limited due to the difficulty of obtaining milligram amounts of purified protein. The cloning of the *ssb* gene into plasmids that cause increase in *ssb* gene expression made the purification of tens of milligrams of SSB protein a routine issue.

In this study, *TaqSSB* protein was cloned into the pET28b vector. Recombinant 6xHis-*TaqSSB* protein was produced in *E. coli BL21 (DE3) PlySE* cells and purified by nickel affinity chromatography. 20 mg of protein was obtained from 1L bacterial culture. It was shown that the *TaqSSB* protein produced and purified at high yield was highly effective in increasing the efficiency and specificity of PCR products. The production of *TaqSSB* protein in high amounts and purity recombinantly in this study may contribute its use as a tool to increase PCR efficiency and specificity.

Acknowledgment

This study was supported by the Turkish Scientific and Technical Research Council (TUBITAK) (TUBITAK-2209A)

Declaration

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article. The authors also declared that this article is original, was prepared in accordance with international publication and research ethics, and ethical committee permission or any special permission is not required.

Author Contributions

All the authors have equally contributed.


References

1. Csako, G., *Present and future of rapid and/or high-throughput methods for nucleic acid testing*. Clinica Chimica Acta, 2006. **363**(1-2): p. 6-31.
2. Ralser, M., et al., *An efficient and economic enhancer mix for PCR*. Biochemical and Biophysical Research Communications, 2006. **347**(3): p. 747-751.
3. Sahdev, S., et al., *Amplification of GC-rich genes by following a combination strategy of primer design, enhancers and modified PCR cycle conditions*. Molecular and Cellular Probes, 2007. **21**(4): p. 303-307.
4. Chen, X.Q., et al., *Betaine improves LA-PCR amplification*. Sheng Wu Gong Cheng Xue Bao, 2004. **20**(5): p. 715-718.
5. Spiess, A.N., N. Mueller, and R. Ivell, *Trehalose is a potent PCR enhancer: lowering of DNA melting temperature and thermal stabilization of taq polymerase by the disaccharide trehalose*. Clinical Chemistry, 2004. **50**(7): p. 1256-1259.
6. Schnoor, M., et al., *Characterization of the synthetic compatible solute homoectoine as a potent PCR enhancer*. Biochemical and Biophysical Research Communications, 2004. **322**(3): p. 867-872.
7. Kinoshita, E., E. Kinoshita-Kikuta, and T. Koike, *A heteroduplex-preferential Tm depressor for the specificity-enhanced DNA polymerase chain reactions*. Analytical Biochemistry, 2005. **337**(1): p. 154-160.
8. Dabrowski, S. and J. Kur, *Cloning, overexpression, and purification of the recombinant His-tagged SSB protein of Escherichia coli and use in polymerase chain reaction amplification*. Protein Expression and Purification, 1999. **16**(1): p. 96-102.
9. Cadman, C.J. and P. McGlynn, *PriA helicase and SSB interact physically and functionally*. Nucleic Acids Research, 2004. **32**(21): p. 6378-6387.
10. Genschel, J., U. Curth, and C. Urbanke, *Interaction of E. coli single-stranded DNA binding protein (SSB) with exonuclease I. The carboxy-terminus of SSB is the recognition site for the nuclease*. Journal of Biological Chemistry, 2000. **381**(3): p. 183-192.
11. Olszewski, M., et al., *Application of SSB-like protein from Thermus aquaticus in multiplex PCR of human Y-STR markers identification*. Molecular and Cellular Probes, 2005. **19**(3): p. 203-205.
12. Dabrowski, S., et al., *Novel thermostable ssDNA-binding proteins from Thermus thermophilus and T. aquaticus-expression and purification*. Protein Expression and Purification, 2002. **26**(1): p. 131-138.
13. Filipkowski, P., A. Duraj-Thatte, and J. Kur, *Novel thermostable single-stranded DNA-binding protein (SSB) from Deinococcus geothermalis*. Archives of Microbiology, 2006. **186**(2): p. 129-137.
14. Filipkowski, P., A. Duraj-Thatte, and J. Kur, *Identification, cloning, expression, and characterization of a highly thermostable single-stranded-DNA-binding protein (SSB) from Deinococcus murrayi*. Protein Expression and Purification, 2007. **53**(1): p. 201-208.
15. Filipkowski, P., M. Koziatek, and J. Kur, *A highly thermostable, homodimeric single-stranded DNA-binding protein from Deinococcus radiopugnans*. Extremophiles, 2006. **10**(6): p. 607-614.
16. Olszewski, M., et al., *Characterization of exceptionally thermostable single-stranded DNA-binding proteins from Thermotoga maritima and Thermotoga neapolitana*. BMC Microbiology, 2010. **10**: 260.
17. Olszewski, M., M. Mickiewicz, and J. Kur, *Two highly thermostable paralogous single-stranded DNA-binding proteins from Thermoanaerobacter tengcongensis*. Archives of Microbiology, 2008. **190**(1): p. 79-87.
18. Bernstein D.A., et al., *Crystal structure of the Deinococcus radiodurans single-stranded DNA-binding protein suggests a mechanism for coping with DNA damage*. Proceedings of the National Academy of Sciences, 2004. **101**(23): p. 8575-8580.
19. Filipkowski P. and J. Kur, *Identification and properties of the Deinococcus grandis and Deinococcus proteolyticus single-stranded DNA binding proteins (SSB)*. Acta Biochimica Polonica, 2007. **54**(1): p. 79-87.
20. Wadsworth R.I. and M.F. White, *Identification and properties of the crenarchaeal single-stranded DNA binding protein from Sulfolobus solfataricus*. Nucleic Acids Research, 2001. **29**(4): p. 914-920.
21. Witte, G., R. Fedorov, and U. Curth, *Biophysical analysis of Thermus aquaticus single-stranded DNA binding protein*. Biophysical Journal, 2008. **94**(6): p. 2269-2279.
22. Jedrzejczak, R., et al., *Structure of the single-stranded DNA-binding protein SSB from Thermus aquaticus*. Acta Crystallographica Section D, 2006. **62**(11): p. 1407-1412.
23. Vallejo, L.F. and U. Rinas, *Strategies for the recovery of active proteins through refolding of bacterial inclusion body proteins*. Microbial Cell Factories, 2004. **3** (1): 11.
24. Kuduğ, H., et al., *Production of red fluorescent protein (mCherry) in an inducible E. coli expression system in a bioreactor, purification and characterization*. International Advanced Researches and Engineering Journal, 2019. **3**(1): p. 20-25.
25. Perales, C., et al., *Enhancement of DNA, cDNA synthesis and fidelity at high temperatures by a dimeric single-stranded DNA-binding protein*. Nucleic Acids Research, 2003. **31**(22): p. 6473-6480.



Research Article

Dose response of gluconic acid doped Fricke gels irradiated with X-rays

Serkan Aktaş^a , Özlem Korkut^{b,*}  and M. Erdem Sağsöz^c 

^aRegional Directorate of Hygiene Laboratory, Erzurum 25200, Turkey

^bFaculty of Eng., Dept. of Chem. Eng., Atatürk University, Erzurum 25240, Turkey

^cFaculty of Medicine, Biophysics Dept, Atatürk University, Erzurum 25240, Turkey

ARTICLE INFO

Article history:

Received 11 July 2020

Revised 12 December 2020

Accepted 21 December 2020

Keywords:

Fricke gels

FT-IR

Gluconic acid

UV

ABSTRACT

Adjusting the dose of radiation which is received by a cancer patient during radiotherapy is very important. The use of dosimetric gels to calculate the applied dose distribution three-dimensionally is a current research topic in radiotherapy. In this study, 16 different Fricke gels including gluconic acid (GA) were produced. These gels were irradiated from 0 to 250 cGy with increments of 50 cGy. MR intensity values and images, UV absorbance values and FT-IR spectra of gels were obtained before and after the irradiation process. The UV absorbance and MR intensity values showed a linear increase in relation to the increase in the applied dose and the amount of ferrous sulfate and GA content in the gels. The oxidation of iron increases as a result of the interaction with hydrogen peroxide which is the product of the irradiation process and GA, and thus the response of the gel to the irradiation process becomes more effective.

© 2021, Advanced Researches and Engineering Journal (IAREJ) and the Author(s).

1. Introduction

The use of radiation in the treatment of cancerous cells is called radiotherapy. It is important to know the dose amount applied to cancer patients. A number of systems were developed to measure the dose of radiation, such as control dosimetry, small area dosimetry, out-of-field low dose dosimetry, in vivo dosimetry, brachytherapy, and auditing of radiotherapy applications [1]. Dosimetric gels can be used for the determination of the dose distribution three dimensionally before the planned treatment. As a result of the interaction of the gel dosimeters with radiation, there are changes in various properties such as optical, magnetic, color, ultrasonic permeability and scattering. Fricke gels are useful as three-dimensional dosimeters because of their ease of production and tissue equivalence [2,3]. The basic change in Fricke gels, which are produced by dissolving iron sulfate solution containing very low amounts of acid and salt in a gel structure, is conversion of Fe^{2+} to Fe^{3+} by irradiation. Fricke gels are generally examined by magnetic resonance imaging (MRI) technique utilizing the different behaviors of ferrous (Fe^{2+}) and ferric (Fe^{3+}) iron in the

magnetic field [4-8]. The main disadvantage of Fricke gel dosimeters is that the ferric ions undergo diffusion after irradiation and this gradually causes the dose pattern to blur. In dosimeters containing xylenol orange (XO), diffusion is slightly reduced and it is used for NMR (nuclear magnetic resonance) analysis especially [9].

There are several studies that analyzed Fricke gels with optical techniques. Gambarini et al. [10] investigated the optical absorbance and magnetic resonance of Fricke XO gel dosimeters prepared with different XO and gelling agents. Gallo et al. [9] examined the optical absorbance spectra (350-750 nm interval) of Fricke XO gel dosimeters loaded with laponite. Optically-enhanced gel dosimeters produced by adding Ethylene diamine tetra acetic acid (EDTA) to the ammonium ferrous sulfate solutions including XO were studied [11]. Fricke XO-gelatin (FXG) and its application as a gel dosimeter in radiotherapy were determined with some parameters such as beam uniformity, optical absorbance and output factor [12]. More recently, Lazzaroni et.al used a ligand having the iminodiacetic and phenol moieties instead of XO in Fricke gels for more accurate and feasible dose evaluation by

* Corresponding author. Tel.: +90-442-2314558; Fax: +90-442-2314910.

E-mail addresses: aktasveserkan@gmail.com (S. Aktaş), ozlemkor@atauni.edu.tr (Ö. Korkut), mesagsoz@atauni.edu.tr (M. E. Sağsöz)

ORCID: 0000-0001-8196-9873 (S. Aktaş), 0000-0002-1427-9183 (Ö. Korkut), 0000-0002-3324-6942 (M. E. Sağsöz)

DOI: 10.35860/iarej.768266

This article is licensed under the CC BY-NC 4.0 International License (<https://creativecommons.org/licenses/by-nc/4.0/>).

optical and MRI analysis [13]. In our study, the dose distribution of X-ray irradiated Fricke gels, including gluconic acid (GA) were examined. Ultra violet (UV) absorbance values and Fourier transform infrared (FT-IR) spectra of gels were obtained before and after the irradiation process. GA, which is found in some fruits and in iron-containing tablets taken as food supplements, was added to Fricke gels in this and a previous [7] study for the first time as a biocompatible material. In this way, the aim was to increase the oxidation of Fe^{2+} to Fe^{3+} , which is the basis for the usage of Fricke gel as dosimeter, by adding reactions of GA. Results of the UV and FT-IR analysis proved these expected chemical reactions occurred.

2. Material and Method

2.1. Materials

Gelatin bovine medical grade (CAS 9000-70-8), D-Gluconic acid (CAS 526-95-4) and Iron(II) sulfate hydrate (CAS 7782-63-0) were purchased from Sigma-Aldrich for production of gels.

2.2. Preparation of Fricke Gels

In this study, firstly 4 different groups of gels were prepared. These groups contained 0, 0.5, 1.0 and 2.0 mM GA ($\text{HOCH}_2(\text{CHOH})_4\text{COOH}$), respectively. Then, $\text{Fe}(\text{NH}_4)_2(\text{SO}_4)_2 \cdot 7\text{H}_2\text{O}$ solutions containing 50 mM H_2SO_4 were added to gels to provide the desired concentrations (0.125, 0.25, 0.5 and 1.0 mM) of Fe^{2+} . Thus, firstly, 16 different types of Fricke gel were prepared. The preparation of gels is given in detail as follows. Deionized water of 100 ml was heated to 85 °C and then 1 g bovine gelatin was added and mixed at 400 rpm for 5 minutes with a magnetic stirrer under air flow at 20 Lh^{-1} . The solution was left to boil at saturation for 20 min and then cooled down to 70 °C. GA was added with the required amount and stirred for one minute. The solution was then removed from the magnetic stirrer and FeSO_4 solutions containing H_2SO_4 were added. After 10-15 seconds of manual mixing, they were poured into spectrophotometry cuvettes with a 1 cm optical path. Each type of gel was poured into 6 cuvettes for 6 different irradiation doses. Thus, finally, 96 different Fricke gel samples were prepared. Prepared gels were kept in a refrigerator at +4 °C until the irradiation process and analyses were carried out. The contents and names of the gels according to the applied dose (0 and a 250 cGy) are given in Table 1. As can be seen from the table, when the gels are named, they were first grouped as A, B, C, D according to their gluconic acid content, then 1, 2, 3, 4 numbers were added according to the amount of iron sulphate and finally applied radiation dose values were written. Other names of 64 samples (irradiated at 50, 100, 150 and 200 cGy) were not shown in the table to avoid taking up much space.

Table 1. Names and contents of the gels

Name of group	FeSO_4 (mM)	Name of gel before irradiation	Name of gel after 250 cGy irradiation
A (Gluconic acid 0 mM)	0.125	A10	A1250
	0.25	A20	A2250
	0.5	A30	A3250
	1.0	A40	A4250
B (Gluconic acid 0.5 mM)	0.125	B10	B1250
	0.25	B20	B2250
	0.5	B30	B3250
	1.0	B40	B4250
C (Gluconic acid 1.0 mM)	0.125	C10	C1250
	0.25	C20	C2250
	0.5	C30	C3250
	1.0	C40	C4250
D (Gluconic acid 2.0 mM)	0.125	D10	D1250
	0.25	D20	D2250
	0.5	D30	D3250
	1.0	D40	D4250

2.3. Irradiation Process, MR Imaging, UV and FT-IR Analyses

After 24 hours, the prepared gels were removed from the refrigerator, allowed to reach room temperature and irradiated with linear accelerator X-rays from 0 to 250 cGy with increments of 50 cGy. The linear accelerator (Siemens, Primus, Erlangen, Germany) generated X-rays at 6 MV and was calibrated using a traceable ionization chamber complying with IAEA TRS-398 standard. The irradiation of gel dosimeters was done with a similar setup to the calibration of the linac system. 6 MV X-rays were used with a dose rate of 600 MU/min at 100 cm SSD and 10 x 10 cm field size under a 0.5 cm water phantom. Total uncertainty was $\pm 2\%$ of dose delivered. Irradiated gel dosimeters were scanned with a 3T MR system (Siemens, Skyra, Germany). Quadrature brain coil was used with standard turbo spin echo protocol with TR = 367 ms and TE = 14 ms. T1 weighted MR images were obtained and MR intensities were acquired from ROIs (region of interest) with software (Siemens, Syngo Via, Germany).

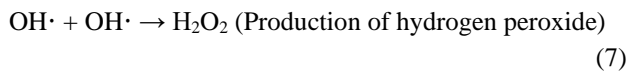
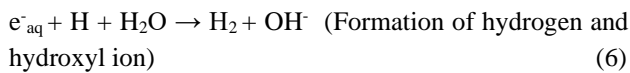
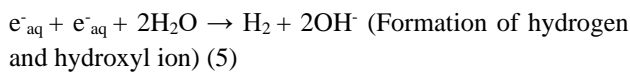
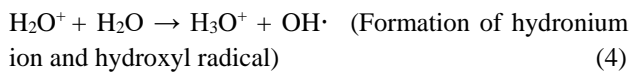
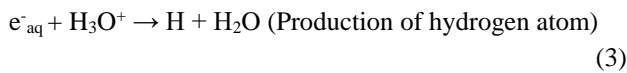
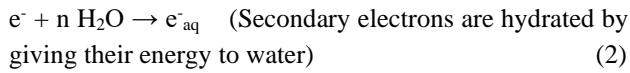
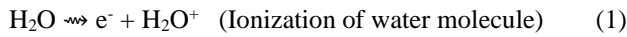
Absorbance values for the Fricke gels were obtained using a Metash 5100 UV / VIS Spectrophotometer at 300 nm wavelength before and after the irradiation process.

FT-IR analysis was performed to monitor changes in the chemical structures of the generated gels. Transmission data were obtained using the Vertex 70 FT-IR device in ATR mode for A40, B40, C40, D40, A4250, B4250, C4250 and D4250 samples.

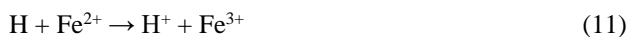
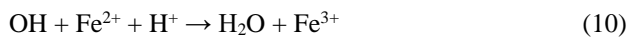
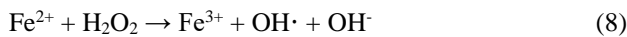
3. Results and Discussion

Before discussing the results of the UV and FT-IR analyses, it is useful to examine the reactions previously described in the literature [14, 15] which are caused by the

irradiation of the gel samples. Since a large part of the solution is composed of water, it is necessary to look at the reactions related to radiolysis of water primarily. Although many reactions take place, the main reactions that occur when water and aqueous solutions are exposed to X-rays are as follows:



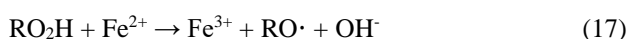
The possible reactions that occur with the irradiation of the Fricke gel systems are as follows. The initial reactions use the products of water radiolysis reactions:



Then, gelatin (RH) which is a macromolecule in the gel system enters reactions with the formed radicals to form new products:



New chain reactions begin with this new radical in the system:



When all these reactions are examined, it appears that many new products are formed in the irradiated gel. The effects of these new species on the results of UV and FT-IR analyses are examined in the following sections.

3.1 UV Measurement Results

The following graphs were obtained according to the results of absorbance at 300 nm before and after the irradiation process.

When Fig. 1 is examined, the absorbance intensities increased with the addition of GA as expected, but the increase in FeSO_4 concentration affected the increase in UV absorbance values more. In particular, gels containing high amounts of FeSO_4 have very high absorbance values. As the amount of applied radiation increases, the intensity of the absorbance increases. These increases in the UV absorbance values were caused by new products (such as) formed during the reactions described above.

3.2 FT-IR Analysis Results of Fricke Gels Non-Irradiated and Irradiated with 250 cGy Dose of X-rays

The wavenumbers of the specific peaks on the FT-IR spectrum for Fricke gels produced in this work and the related bond structures are given in Table 2. If the spectrum obtained in this study is compared with the FT-IR profile of pure gelatin given by Hermanto et. al. [16] and Hossan et.al. [17], the peak values given in Table 2 at 3294, 3078, 2941, 1635, 1539 and 868 cm^{-1} belong to gelatin. These values are very similar to the 3222, 3100, 2947, 1637 and 1542 cm^{-1} peaks, respectively, given by Hossan et al. [17]. In the literature, Fe-O and Fe-OH peaks occur in the fingerprint region of 400 - 1500 cm^{-1} .

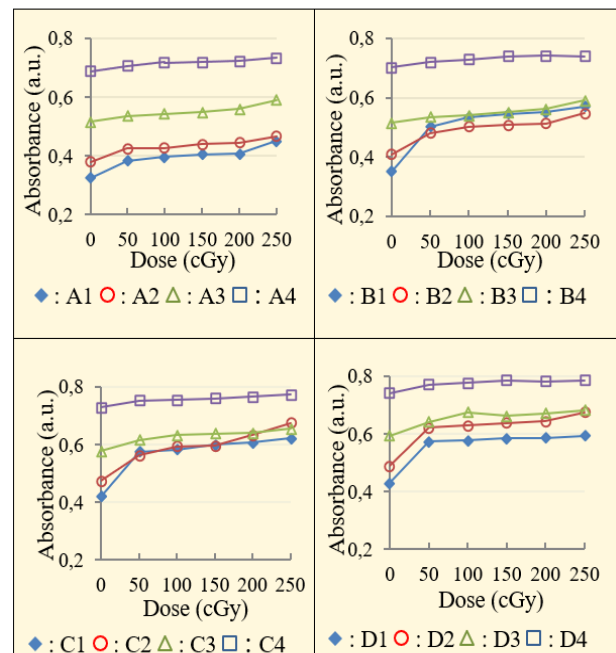


Figure 1. Dose (cGy) and absorbance graph of the gels before and after the irradiation process.

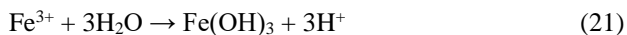
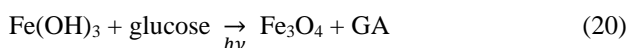
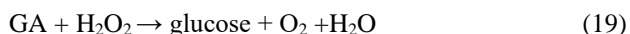
Table 2. The bond structures and wavenumbers for the main peaks in the FT-IR graphs presented in Figure 2 and Figure 3

Wavenumber (cm ⁻¹)	3294	3078	2941	1635	1539	1167	1045	868	577
Bond structure	N-H stretching, -OH stretching	C-H stretching	C-H stretching	C=O stretching	C-N-H bending	C-O stretching	C-O	N-H out of plane bending	Fe-O stretching

Table 3. Slope of dose response lines and R² values for each gel

	Name of gels							
	A10-A1250	A20-A2250	A30-A3250	A40-A4250	B10-B1250	B20-B2250	B30-B3250	B40-B4250
Slope	0.663	0.977	1.227	1.958	0.864	2.250	2.034	2.486
R ²	0.948	0.968	0.963	0.876	0.861	0.967	0.987	0.962
	Name of gels							
	C10-C1250	C20-C2250	C30-C3250	C40-C4250	D10-D4250	D20-D2250	D30-D3250	D40-D4250
Slope	2.217	2.585	2.044	2.051	1.138	1.811	2.138	1.845
R ²	0.992	0.902	0.883	0.885	0.804	0.941	0.930	0.934

Sun et al. [18] studied the structure of nano-magnetite in the presence of GA and Gündüz and Bayrak [19] used nanoscale iron for an adsorption process. They indicated that the peaks showing the Fe-O bond formed at 579 cm⁻¹. Xiao et al. [20] found that magnetite (Fe₃O₄) has strong absorbance at 570 cm⁻¹. Similarly, Ercan et.al. showed the absorption bands for their synthesized iron oxide nanoparticles at 599.83 and 475 cm⁻¹ as related for Fe-O stretches of Fe₃O₄ and Fe₂O₃ [21]. So, the peak at 577 cm⁻¹ shows the presence of Fe-O bond and Fe₃O₄ structure in Figures 2 and 3. According to Chen et al. [22], FT-IR/ATR analysis of glucose solutions showed absorbance peaks around 3300 and 1650 cm⁻¹. C-O bands of the spectrum at 1167 and 1045 cm⁻¹ are evidence of the presence of acid. When all these data are examined, the following reactions can be suggested as causing the changes in the spectrum:

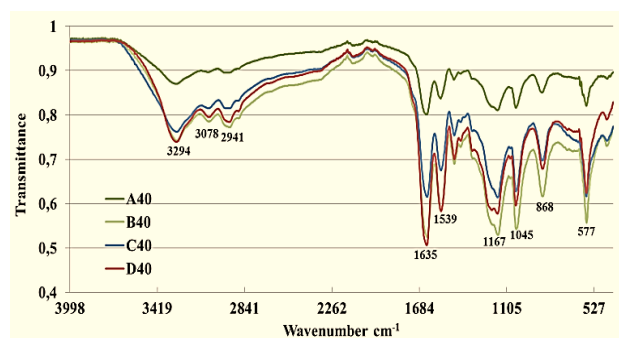
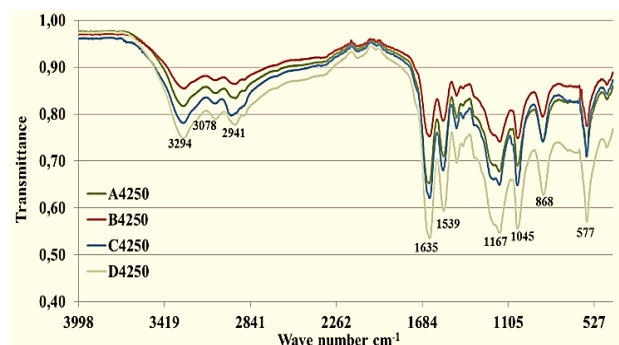


Considering these suggested reactions; it can be said that the oxidation of iron increases as a result of the interaction of hydrogen peroxide, which is the product of the irradiation process, and GA and thus the response of the gel to the irradiation process becomes more effective. The increases in the UV and MR intensity values according to the GA amount confirm this prediction.

As the amount of Fe²⁺ concentration increases, the MR intensity values increase linearly. Similarly, the gels with high GA content have a better dose response trend. In general, the dose response relationship of gels containing 0.5 and 1 mM GA appears to be better.

3.3. MR Intensity Values and Images of the Gels

The MR intensity values were plotted against the irradiation dose. The slope of the straight lines (proportional to the radiation sensitivity of gels) and R² values are shown in Table 3. Figure 4 shows the MR images of the gels using an MR sequence with TR 367 ms and TE 14 ms. As can be seen in this figure, as the amount of iron and GA in the gels increases, brighter areas appear on MR images.

Figure 2. FT-IR graphs of Fricke gels including 1.0 mM FeSO₄ and different amounts of GA before the irradiation processFigure 3. FT-IR graphs of Fricke gels including 1.0 mM FeSO₄ and different amounts of GA after the irradiation process with 250 cGy dose of X-rays

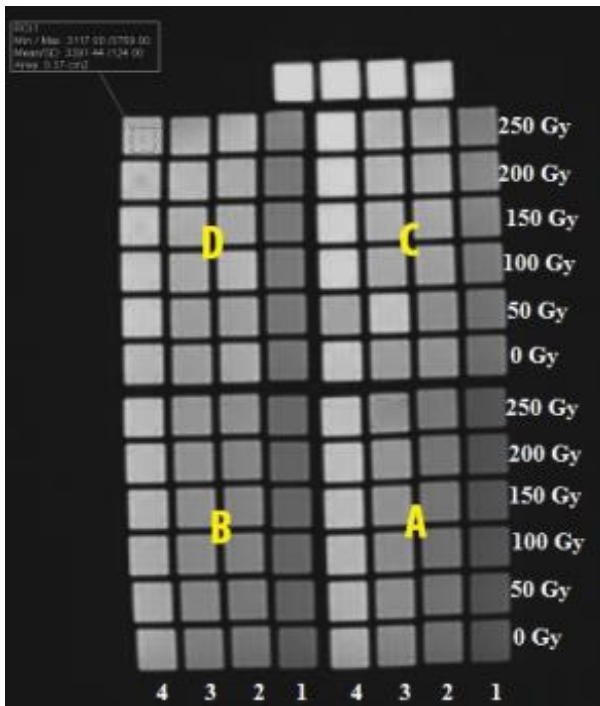


Figure 4. MR Images of the gels containing GA (A: 0 mM, B: 0.5 mM, C: 1 mM, D: 2 mM) and FeSO₄ solutions (1: 0.125 mM, 2: 0.25 mM, 3: 0.5 mM, 4: 1.0 mM)

4. Conclusions

In this study, Fricke gels containing FeSO₄ and GA with different concentrations were produced and then irradiated up to 250 cGy with increments of 50 cGy. UV and FT-IR analyses were performed on the samples before and after irradiation. The UV absorbance values of the gels containing GA increased linearly with the applied radiation dose. Similarly, the MR intensity values showed a linear increase in correlation to the increase in the applied dose and the amount of ferrous sulfate and GA content of the gels. These increases are related to the oxidation of iron depending on chemical reactions where reactants are GA and products of irradiation process.

Declaration

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article. The authors also declared that this article is original, was prepared in accordance with international publication and research ethics, and ethical committee permission or any special permission is not required.

Author Contributions

All authors conceived the study together. Ö. Korkut developed methodology and analysis, S. Aktaş performed experiments and M.E. Sağsöz supervised and improved the study.

References

1. Kron, T., Lehmann, J., Greer, P.B., *Dosimetry of ionising radiation in modern radiation oncology*. Physics in Medicine & Biology, 2016. **61**(14): R167.
2. Keal, P., Baldock C., *A theoretical study of the radiological properties and water equivalence of Fricke and polymer gels used for radiation dosimetry*. Australasian Physical & Engineering Sciences in Medicine, 1999. **22**(3): p. 85-91.
3. Gambarini, G., Carrara, M., Marrani, M., Pirola, L., Tomatis, S., Valente, M., Vanossi, E., *Optical analysis of gel dosimeters: Comparison of Fricke and normoxic polymer gels*. Nuclear Instruments and Methods in Physics Research, 2007. **263**: p. 191-195.
4. Galante, A.M.S., Cervantes, H.J., Cavinato, C.C., Campos, L.L., Rabbani, S.R., *MRI study of radiation effect on Fricke gel solutions*. Radiation Measurement, 2008. **43**(2-6): p. 550-553.
5. Ibbott, G.S., Roed, Y., Lee, H., Alqathami, M., Wang, J., Pinsky, L., Blencowe, A., *Gel dosimetry enables volumetric evaluation of dose distributions from an MR-guided linac*. AIP Conference Proceedings, 2016 Jun 17, Article number 040002.
6. Soliman, Y.S., El Gohary, M.I., Abdel Gawad, M.H., Amin, E.A., Desouky, O.S., *Fricke gel dosimeter as a tool in quality assurance of the radiotherapy treatment plans*. Applied Radiation and Isotopes, 2017. **120**: p. 126-132.
7. Aktaş, S., Korkut, Ö., Sağsöz, M.E., *Dosimetric fricke gel systems improved with CaCl₂ and gluconic acid*. International Advanced Researches and Engineering Journal, 2018, **2**(2):143-146.
8. De Deene, Y., Gel – based radiation dosimetry using quantitative MRI, Chapter 9, NMR and MRI of Gels, 2020, Editor: Yves De Deene p. 275-357.
9. Gallo, S., Cremonesia, L., Gambarini G., Ianni L., Lenardi, C., Argenti, S., Bette, D., Gargano, M., Ludwig, N., Veronese, I. *Study of the effect of laponite on Fricke xylenol orange gel dosimeter by optical techniques*. Sensors and Actuators B, 2018. **272**: p. 618-625.
10. Gambarini, G., Veronese, I., Bettinelli, L., Felisi, M., Gargano, M., Ludwig, N., Lenardi, C., Carrara M., Collura G., Gallo S. et al. *Study of optical absorbance and MR relaxation of Fricke xylenol orange gel dosimeters*. Radiation Measurement, 2017. **106**: p. 622-627.
11. Babu, S.E.S., Peace, B.S.T, Rafic, K.M, Raj E.W.M, Christopher J.S, Ravindran B.P. *Escalation of optical transmittance and determination of diffusion coefficient in low-bloom strength gelatin-based Fricke gel dosimeters*. Radiation Physics and Chemistry, 2019. **156**: p. 300–306.
12. Abdelgawad, M.H, Soliman, Y.S, ElGohry, M.I, Eldib, A.A., Ma C.-M.C., Desouky, O., *Measurements of radiotherapy dosimetric parameters using Fricke gel dosimeter*. Biomedical Physics & Engineering Express, 2017. **3** (2): p. 025021.
13. Lazzaronia, S., Liosic, G.M, Marjanic, M., Dondia D., *An innovative Fe³⁺ selective ligand for Fricke-gel dosimeter*, Radiation Physics and Chemistry, 2020, **171**: 108733.
14. Keller, B.M. *Characterization of the NMR-based Fricke-gelatin radiation dosimeter*. Master of Science in Medical Physics. 1994. Montreal: McGill University.
15. Spinks, J.W.T, Woods, R.J. *An introduction to radiation chemistry*. 1990. New York (NY): Wiley; ISBN 0-471-61403-3.
16. Hermanto, S., Sumarlin, L.O, Fatimah W., *Differentiation of bovine and porcine gelatin based on spectroscopic and*

electrophoretic analysis. Journal of Food and Pharmaceutical Sciences, 2013. **1**: p. 68-73.

17. Hossan, Md.J, Gafur, M.A, Kadir Md.R, Karim M.M. *Preparation and characterization of gelatin- hydroxyapatite composite for bone tissue engineering*. International Journal of Engineering & Technology Sciences, 2014, **14**(01): p.24.
18. Sun, X, Zheng, C, Zhang, F, Yang Y, Wu G, Yu A, Guan N. *Size-controlled synthesis of magnetite (Fe_3O_4) nanoparticles coated with glucose and GA from a single Fe(III) precursor by a sucrose bifunctional hydrothermal method*. Journal of Physics and Chemistry C, 2009. **113**: p. 16002–16008.
19. Gündüz F, Bayrak B., *Synthesis and performance of pomegranate peel-supported zero-valent iron nanoparticles for adsorption of malachite green*. Desalination Water Treatment, 2018. **110**: p. 180–192.
20. Xiao W, Jones AM, Collins RN, Bligh MW, Waite TD. *Use of fourier transform infrared spectroscopy to examine the Fe(II)-Catalyzed transformation of ferrihydrite*. Talanta, 2017, **175**: p. 30–37.
21. Ercan, G., Uzunoğlu, D., Ergüt, M., Özer, A., *Biosynthesis and characterization of iron oxide nanoparticles from Enteromorpha spp. extract: determination of adsorbent properties for copper (II) ions*. International Advanced Researches and Engineering Journal. 2019, **3**(1): p. 65-74.
22. Chen, J., Wu, L., Pan, T., Xie, J., Chen, H. *A quantification method of glucose in aqueous solution by FTIR/ATR spectroscopy*. Seventh International Conference on Fuzzy Systems and Knowledge Discovery (FSKD 2010), 2010, p. 2159-2163.



Research Article

GMDH-type neural network-based monthly electricity demand forecasting of Turkey

Ali Volkan Akkaya ^{a,*} 

^aYildiz Technical University, Department of Mechanical Engineering, 34349 Besiktas, Turkey

ARTICLE INFO

Article history:

Received 09 July 2020

Revised 25 October 2020

Accepted 02 November 2020

Keywords:

Electricity demand
Forecasting
GMDH neural network
Modeling

ABSTRACT

In this study, it was intended to develop an accurate forecasting model for the monthly electricity demand of Turkey in the medium-term. For this purpose, the Group Method of Data Handling (GMDH)-type Neural Network (NN) approach was utilized to structure a nonlinear time-series based forecasting model. A large dataset containing monthly electricity demand was considered for the period of 2003-2018. The developed model was tested in the period of 2019/01-2019/11 in order to specify the generalization ability. The test results showed that the developed model was very close to actual values. The obtained test performances were 2.10 % for mean absolute percentage error (MAPE), 2.36 % for root mean square percentage error (RMSPE) and 0.869 for coefficient of determination (R^2). In addition, results of the developed GMDH-type NN model were compared to the forecasting results of a literature study. The comparison revealed that GMDH-type NN was a better approach for forecasting the monthly electricity demand for Turkey. Finally, the developed model was utilized to forecast monthly electricity demand in the period of 2019/12-2020/12.

© 2021, Advanced Researches and Engineering Journal (IAREJ) and the Author(s).

1. Introduction

Currently, the technical, economic and social development of any nation state in the world depends largely on the utilization of electrical energy [1]. Electricity is required for almost all types of human actions such as manufacturing, agriculture, housing, heating, lighting, and transportation. Moreover, the electricity consumption per capita that is a direct measure of the standard of living is used as a descriptive indicator to determine the development stage of countries [2]. Furthermore, based on primarily growing population and desired welfare and living conditions, the use of electricity increases from year to year [3]. From this point of view, future electricity demand in a country should be accurately forecasted and carefully planned in order to maintain the demand-supply balance by making the required investments on time [4].

In this aspect, Turkey is located at a strategic geographical location connecting from Middle-East and Asia regions, where conventional energy sources are

abundant, to Europe where energy consumption is significantly high [5]. However, Turkey has inadequate oil and natural gas reserves. They are 44.3 million tons for oil and 6.2 million m^3 for natural gas. Hydroelectric, low-quality lignite coal, and wind are the main domestic energy resources of Turkey. Therefore, it is extremely dependent on foreign fossil fuels to meet the growing electricity demand [5]. Electricity demand increases day by day in Turkey regarding economic and population growth [6,7]. The country's total electricity generation increased from 23.275 TWh in 1980 to 304.800 TWh in 2018. At the same time, with the technological progress (smart grids), electricity demand forecasting becomes more critical for grid operators, market participants, regulators, and ministers in terms of planning and operational decisions. It is obvious that precise electricity demand forecasting is crucial for the protecting the finite resources, but it is still a problematic issue because of the stochastic and ambiguous characteristics [8].

In the related literature, a number of studies concerning electricity demand forecasting have been carried out based

* Corresponding author. Tel.: +90 212 3832786

E-mail addresses: aakkaya@yildiz.edu.tr (A. V. Akkaya)

ORCID: 0000-0001-7189-592X

DOI: 10.35860/iarej.766762

This article is licensed under the CC BY-NC 4.0 International License (<https://creativecommons.org/licenses/by-nc/4.0/>).

on different methods over the years. Generally, load/demand forecasting methods can be separated to two leading categories: i) Conventional Statistical, and ii) Artificial Intelligence (AI) methods [8,9]. The conventional methods contain multiple linear regression, similar-day techniques, exponential smoothing models, semi-parametric additive models and time series modeling including seasonal autoregressive integrated moving average (SARIMA), autoregressive integrated moving average (ARIMA), autoregressive moving average (ARMA), etc. [10-16]. On the other hand, the artificial intelligent methods consist of artificial neural networks (ANNs), fuzzy logic, genetic algorithm, grey prediction, adaptive neuro fuzzy system, expert systems, support vector machines, and hybrid methods [2,3,17-22]. Not every method is appropriate for all forecasting problems because there is not an agreement on the best method. However, contemporary studies have shown the power of ANNs in forecasting electricity demand [20, 23-26]. However, the challenging point in designing ANNs is the choice of the finest network architecture that can provide the best results. This is a great task since the architecture of the network includes many important points, for example, the neuron number in different layers, layer number, and number of inputs and outputs [24]. When the Group Method of Data Handling (GMDH) is applied in the building of ANNs, the main problems mentioned above can be overcome because it has the ability to generate the network automatically [27]. In this study, the GMDH approach was used as one of the Neural Networks (NNs), which was named GMDH-type NN, and it is explained in detail in the next section.

The demand forecasting can be grouped namely with regard to the time horizon. It can be typically divided into three periods. First group is long term specified from a year to a decade. Second group is medium-term quantified from a week to a year. Third group is short-term measured from 1 h to 1 week. Acquiring important knowledge for the cost-effective and safe operation of the power generation systems is aimed at short-term forecasting. The long-term forecasting is applied for investment planning and decisions. Instead, the medium-term forecasting contributes on meeting load conditions, fuel purchasing plans, outage & maintenance arrangement, organization of load dispatch and bill payment, cost-effective operation of the electricity generation system, and improved agreement debates in electricity trading [8,9,28]. In this study, medium-term forecasting is considered with placing special emphasis on the monthly electricity demand of Turkey. It is seen in the associated literature that although a number of studies related to the estimating Turkey's electricity demand, most of these studies have been conducted to forecast annual demand for long term [2,4,5,7,10,11,18,19,22]. It is clear that there is a lack of

studies regarding the medium-term electricity demand of Turkey. Only, a few studies are related to the monthly electricity demand forecast even though it has significant roles in the planning and marketing of electricity generation systems. For instance, İlseven and Göl [8] applied multivariate adaptive regression splines technique to forecast monthly electricity demand of Turkey. They indicated that their model achieved successful results in validation step by testing error and their model showed steady forecasting performance. They concluded their study by forecasting monthly electricity demand for the years of 2017-2019. Hamzacebi et al. [26] proposed four different seasonal artificial neural network models and chose the best one to predict Turkey's monthly electricity demand. They finalized their study by implementing the best ANN structure to estimate monthly electricity demand for the years from 2015 to 2018.

In this study, the medium-term based Turkey's electricity demand was forecasted by the GMDH-type NN approach. The core novelty of this study is to implement the first time the GMDH-type NN to Turkey's monthly electricity demand forecasting. For this purpose, the GMDH-type NN model was developed using the monthly dataset of the years 2003-2018 and tested in the period of 2019/01-2019/11 to reach highly accurate predictions. The graphical and statistical analyses were used to show model success and consistency. Besides the forecasting outcomes of a related literature study were compared with the results of the suggested GMDH-type NN model. Lastly, the developed model was applied in the period of 2019/12-2020/12 to estimate monthly electricity demand.

2. GMDH-type Neural Network

GMDH can be considered as a self-organizing ANN which was developed to model multi-variable and non-linear complex systems based on the relationships between inputs and outputs by Alexey G. Ivakhnenko in 1971 [29]. GMDH-type NN, fundamentally a feed-forward and multi-layered neural network, has been implemented to several engineering problems [30- 33].

The training procedure of the GMDH-type Neural Network is based on evolutionary mechanism, unlike the conventional neural network [34]. The GMDH algorithm is represented by a series of neurons in which different pairs of neurons are coupled via a quadratic polynomial and generate new neurons in the subsequent layer. The key input variables, number of layers, number of active neurons, and neurons in hidden layers are automatically constructed by the self-organization of the neural network [33]. Employing an iterative mechanism, the model architecture is adapted to produce the maximum accuracy without any overfitting in data prediction. In other words, to optimize the network, the links among neurons of the related network does not stay constant, on the contrary, they are chosen during the training

process [35,36]. Fig.1 illustrates the general structure and configuration process of GMDH-type NN.

This network identifies the approximate function of \hat{f} through the predicted output of \hat{y} with the minimum error. Quadratic polynomials resulted from all neurons for a set of multi inputs $X = (x_{i1}, x_{i2}, x_{i3}, \dots, x_{in})$ are combined and the predicted output is compared to the actual single output of y [33]. Therefore, the actual results for the experimentally measured M data consisting of n inputs and single-output can be stated as following [31]:

$$y_i = f(x_{i1}, x_{i2}, x_{i3}, \dots, x_{in}) \quad (i=1, 2, 3, \dots, M) \quad (1)$$

To estimate the value of \hat{y} for the input vector of X , a GMDH-type NN is developed as below:

$$\hat{y}_i = \hat{f}(x_{i1}, x_{i2}, x_{i3}, \dots, x_{in}) \quad (i=1, 2, 3, \dots, M) \quad (2)$$

The GMDH-type NN have to minimize the squared error between the estimated and the actual values given as following:

$$\sum_{i=1}^M (\hat{y}_i - y_i)^2 \rightarrow \min \quad (3)$$

The mathematical description between the output and input variables is formulated with utilizing a complex discrete procedure of the functional series recognized as the Kolmogorov–Gabor polynomial [29,32], as given in Eq.(4):

$$y = a_0 + \sum_{i=1}^n a_i x_i + \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j + \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n a_{ijk} x_i x_j x_k + \dots \quad (4)$$

Where n is the number of input variables, (x_1, x_2, \dots, x_n) are the input parameters, and (a_0, a_1, \dots, a_n) are the coefficients. The quadratic and bivariate structure of this polynomial expression is employed usually by total algebraic arrangement [36] given as following:

$$\hat{y}_i = G(x_i, x_j) = a_0 + a_1 x_i + a_2 x_j + a_3 x_i^2 + a_4 x_j^2 + a_5 x_i x_j \quad (5)$$

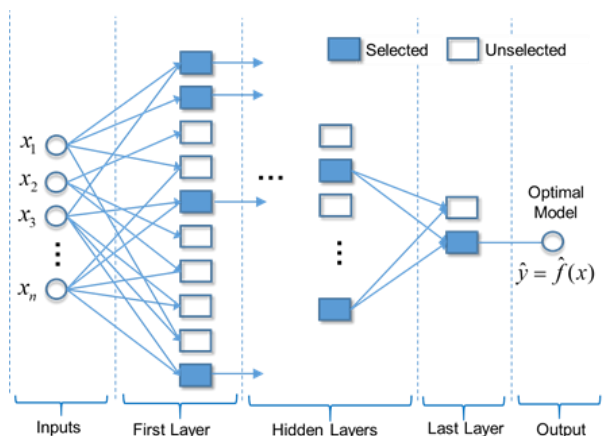


Figure 1. Configuration process of GMDH-type neural network

In order to produce the overall mathematical relationship of input and output variables presented in Eq.(4), such partial quadratic structure is implemented in a opposite route throughout a network of joint neurons [33]. To determine the unknown coefficients of a_i in Eq.(5), the regression technique is performed. Thus, for each pair of input variables $(x_i$ and $x_j)$, the minimum difference between the estimated values (\hat{y}) and the actual output (y) is provided [31]. By means of Eq.(5), a group of polynomials is established. The least squares technique is implemented to determine the unidentified coefficients of the mentioned polynomials. For the constants for each quadratic function G_i , the coefficients are acquired to minimize the total neuron error in order to achieve optimal fit of the inputs, as follows:

$$E = \frac{\sum_{i=1}^M (y_i - G_i)^2}{M} \rightarrow \min \quad (6)$$

Taken into consideration the GMDH procedure, two independent possibility parameters are made available by the whole set of n chosen input parameters. Similarly, the stated technique, $n(n-1)/2$ neurons are going to set through first hidden layer computed from $\{(y_i, x_{ip}, x_{iq}); (i=1, 2, \dots, M)\}$ on behalf of different $p, q \in \{1, 2, \dots, n\}$. Hence, M data triples $\{(y_i, x_{ip}, x_{iq}); (i=1, 2, \dots, M)\}$ are determined through the $p, q \in \{1, 2, \dots, n\}$ from the observation as follows [31,35]:

$$\begin{bmatrix} x_{1p} & x_{1q} & y_1 \\ x_{2p} & x_{2q} & y_2 \\ x_{3p} & x_{3q} & y_3 \end{bmatrix} \quad (7)$$

The second-order form of the expressed function (Eq. 5) is utilized for each M triple row. These equations expressed in the following matrix form:

$$Aa = Y \quad (8)$$

$$a = \{a_0, a_1, a_2, a_3, a_4, a_5\} \quad (9)$$

$$Y = \{y_1, y_2, y_3, \dots, y_M\}^T \quad (10)$$

where a denotes the vector of unknown coefficients presented in Eq.(5) and Y refers to the outputs' vector. Thus, the formulated correlation is given as following:

$$A = \begin{bmatrix} 1 & x_{1p} & x_{1q} & x_{1p}^2 & x_{1q}^2 & x_{1p}x_{1q} \\ 1 & x_{2p} & x_{2q} & x_{2p}^2 & x_{2q}^2 & x_{2p}x_{2q} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{Mp} & x_{Mq} & x_{Mp}^2 & x_{Mq}^2 & x_{Mp}x_{Mq} \end{bmatrix} \quad (11)$$

With examining the multiple-regression outcome, the least square technique is completed by Eq. (12):

$$a = (A^T A)^{-1} A^T \quad (12)$$

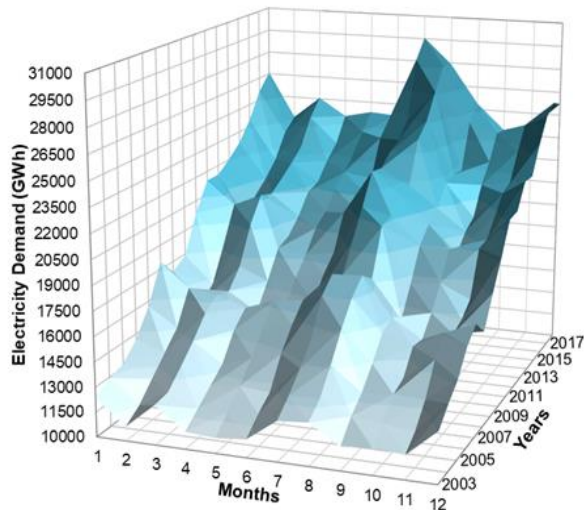


Figure 2. Monthly gross electricity demand of Turkey between 2003 and 2018

The most precise values of the coefficients given in Eq.(5) for M triple datasets of inputs are determined by the Eq.(12). In this architecture, the training algorithm is initiated by the input layer which is only a starting layer and does not comprise any process. In the first layer, based on all probable arrangements of input variables, the neuron candidates are produced. Then, the coefficients of Eq.(5) are determined for each respective neuron. By means of both the determined coefficients and input variables for each neuron, the desired output is estimated. Some neurons with well prediction ability are chosen according to external criteria and fed forward to train the next layer. The neurons which do not cover the considered condition are disregarded from the structure of the network. The outcomes achieved with the chosen neurons turn into the inputs for the subsequent layer. This procedure carry on up to the final layer. In the final layer, only one neuron is chosen. The achieved result in the final layer is the estimated values.

3. Model Development

The main aim of the study is to estimate Turkey's monthly electricity demand by employing GMDH-type NN method. Therefore, in this division, methodical steps of model development comprising of data collection, model configuration, and determination of performance criteria are given to achieve a successful forecasting result. Brief information about these steps is introduced in the subsequent subsections

3.1 Data Collection

Monthly gross electricity demand data for Turkey was obtained from the Turkish Electricity Transmission Company [37]. The data cover a period of monthly electricity demand values from January 2003 to November 2019. The data up to the year of 2019 are given as graphically in the

Fig.2. The electricity demand has been increasing over the years in Turkey. In the given period, electricity demand rose by about 6% annually. There are seasonal behavior and monthly trend in data. Seasonally, whereas the demand for electricity is high in winter and summer months and reaches a peak usually in July or August, and it is low in autumn and spring months and gets the lowest point generally in February or October. This kind of seasonality behavior explains the connection between the climate situations in different seasons and electricity demand.

The collected data were analyzed, interpreted and apportioned into the model with a basic routine. The data were separated into two dataset. They are named as model development and testing parts. The model development dataset included monthly data in the period of 2003/01 – 2018/12, while the testing set comprised the period of 2019/01- 2019/11. The model development dataset was used for the configuration of the GMDH-type NN model. On the other hand, the testing set was used to measure the capability of generalization of the developed model.

3.2 Model Setup

The model was developed using the software package named as GMDH Shell 3.8.9. In this section, the procedures and steps for developing a forecast model by using GMDH Shell are given systematically. The monthly electricity demand data including time-stamps information were compiled using the MS Excel spreadsheet. The excel file was introduced to the modeling media in CSV/XLS/XLSX structure. Model development and testing datasets were adjusted as defined in Subsection 3.1. The dataset was preprocessed to transform data in accordance with configured modeling conditions in the module of data explorer. The input and target (output) variables were specified by several transformations such as elementary functions (e.g. sin, cos, cube), time series (e.g. lags, weighting, moving average), date/time extracting information from timestamps (e.g. month, year), weighted instances setting custom weights for target instances and so on.

Then, the solver module was used to produce predictive models for the target variable. The first thing to do was reorder observation, which was utilized to accomplish unvarying statistical features of training and testing samples as well as to make these parts similarly explanatory. The pseudorandom option was selected for the reorder observation process. The k-fold method was used to model the validation strategy and sorting out. The k-fold validation separates dataset into k sections, trains a model k times utilizing k-1 sections, and measures model performance each time utilizing a fresh left behind portion. As a final point, for comparing model, residuals acquired from entire testing parts were included and utilized.

Table 1. Main properties of the developed model in GMDH solver section

Properties	Selected option
Reorder observation	Pseudo-random
Validation strategy	k-fold
Validation criterion	RMSE
Variable ranking	No
Core algorithm	GMDH neural network
Neuron function	Polynomial and quadric polynomial

The root mean square error (RMSE) was utilized for a model selection measure, which selected the models with the lowest RMSE calculated for the testing sample. Polynomial neural networks of GMDH-type were used as the statistical learning algorithm. A different number of input variables can be permitted for a neuron. Using two inputs for any neuron is significantly effective way. Apart from that, the calculation charge can be excessively difficult. A polynomial and quadratic polynomial can be preferred as an interior function for neurons. To enhance the general estimation ability of the model, each neuron can eliminate some of the function parts. Specifying the maximum number of layers controls the higher boundary of the NN layers generated through the GMDH algorithm. The number of neurons that will be included to the set of inputs in each subsequent layer is determined by the initial layer width. The selected main properties of the model are given in Table 1. The other specified properties are presented in the Results and Discussion section.

3.3 Performance Criteria

It is expected from a model that the estimated values should be close to actual values as much as possible. Therefore, the aim is to minimize the error, in other words, to diminish the difference between actual and estimated values. In this work, the prediction performance results of the time series based GMDH-type NN model for both the model development and testing stages were evaluated by using some performance criteria such as maximum negative error (MNE), maximum negative percentage error (MNPE), maximum positive error (MPE), maximum positive percentage error (MPPE), mean absolute error (MAE), mean absolute percentage error (MAPE), root mean square error (RMSE), root mean square percentage error (RMSPE), correlation (R), and coefficient of determination (R^2). Among them, R and R^2 values close to unity indicate a satisfactory result. Therefore, the highest values are anticipated for these criteria. In contrast, for other mentioned criteria, low value or value close to zero is desired for the forecasting model.

4. Results and Discussion

4.1 Results for Model Development and Testing Stages

Apart from the given information about model development in Section 3, some parameters of the GMDH-

type NN model were set by the trial-error method to achieve the best model results. For this purpose, a number of experiments were executed by the changing number of folds in the validation strategy, neuron function, the maximum number of layers, and initial layer width. The best results were obtained for 2 folds, polynomial neuron function, the maximum layer of 500, and 3 initial layer width.

For these arrangements, the prediction results and residuals from the actual values are given in Fig.3. It is seen that predictions can capture seasonal changes in the years and have the ability to track the actual values. The performance criteria obtained for the model development stage are given in Table 2 and they support the judgment. For example, the maximum positive and negative percentage errors are far below 10 %. MAPE and RMSPE are approximately close to 2 % while R and R^2 values are above 0.95.

The results of the testing stage are depicted in Fig 4. It was understood from this figure that the model prediction results virtually compatible through the actual values. As seen in Table 2, although MNPE, MPPE, MAPE, and RMSPE values are better than the model development stage, R and R^2 values are slightly lower. Based on these performance results, the adequacy scale of the developed model can be evaluated as 'well level'.

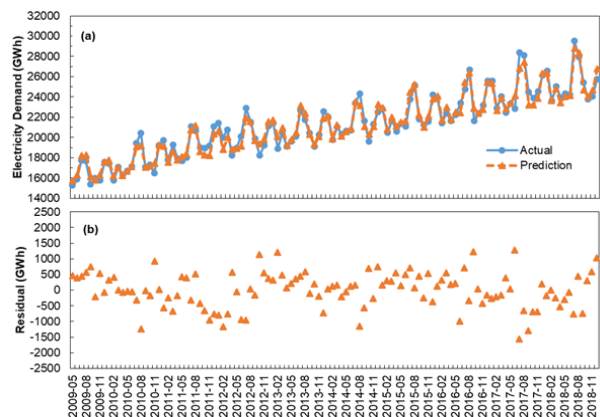


Figure 3. Model development results: (a) actual and prediction values, (b) residuals

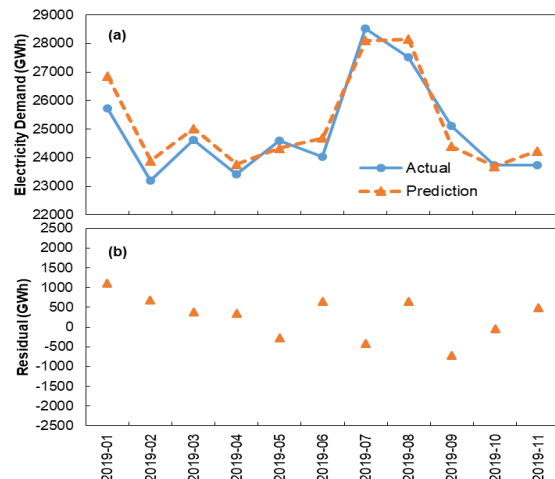


Figure 4. Testing results: (a) actual and prediction values, (b) residuals

Table 2. Performance results of model development and testing stages

Performance criteria	Symbol	Model development	Testing
Maximum negative error	MNE	-1563.33	-725.4
Maximum negative percentage error	MNPE	-6.10 %	-2.89 %
Maximum positive error	MPE	1285.76	1117.4
Maximum positive percentage error	MPPE	6.26 %	4.34 %
Mean absolute error	MAE	449.35	526
Mean absolute percentage error	MAPE	2.12 %	2.10 %
Root mean square error	RMSE	567.46	592.04
Root mean square percentage error	RMSPE	2.67 %	2.36 %
Correlation	R	0.983	0.946
Coefficient of determination	R ²	0.965	0.869

4.2 Comparison with A Study in The Literature

To show the generalization ability of the considered approach, the developed model was compared with a literature study performed by Hamzacebi et al. [26]. In the literature study, the seasonal ANN method was utilized to estimate the monthly electricity demand of Turkey for the period of 2015/01 – 2018/12 by using the dataset covering the period of 2002/01-2014/12. For comparison purposes, the GMDH-type model used in this study was reconfigured for the same forecasting period. The model was developed using the dataset covering the period of 2003/01-2014/12 to forecast the same period examined in the literature study. The number of folds, neuron function and initial layer width were set to 7, quadratic polynomial, and 3, respectively. Based on these arrangements, the GMDH-type NN model results were obtained. Fig.5 graphically shows the comparison of this study and the literature study in terms of the monthly electricity demand forecast. It can be seen in Fig.5(a) that the results of the GMDH-type NN model are superior to that of the literature study in the investigated period. Although the literature study can provide satisfactory results for the period of 2015/01-2015/12, its forecast performance is worsening in the later periods. The same outcomes can be observed when Fig.5(b) are examined. Although residuals of the GMDH-type NN model stay almost the same range throughout all periods, negative residuals of literature study are increasing. In addition, performance criteria of these two studies are compared in Table 3. Almost all criteria point out that the GMDH-type NN based model has a good capability to forecast the monthly electricity demand.

4.3 Forecasting for the Period of 2019/12 – 2020/12

The performed analyses had displayed that the GMDH-type NN model was the robust forecasting tool for the

monthly electricity demand of Turkey. Therefore, the forecasting was executed by this model for the period of 2019/12- 2020/12. The forecasting results achieved by the GMDH-type NN model is depicted in Fig.6. According to this, in the investigated period, the trend in electricity demand is relatively the same for the year 2019. The confidence band indicating the difference of average forecast value from upper or lower one is 1134.92 GWh. Accordingly, it can be said that the peak electricity demand would be 28382 GWh in August of 2020, whereas the bottommost one could be 22995 GWh in November of 2020.

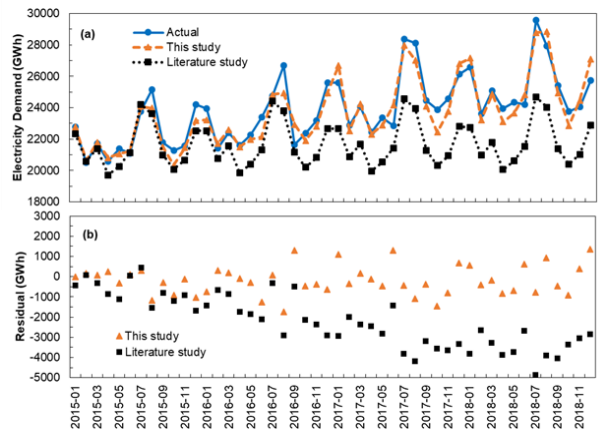


Figure 5. Comparison with the literature study: (a) actual and prediction values, (b) residuals

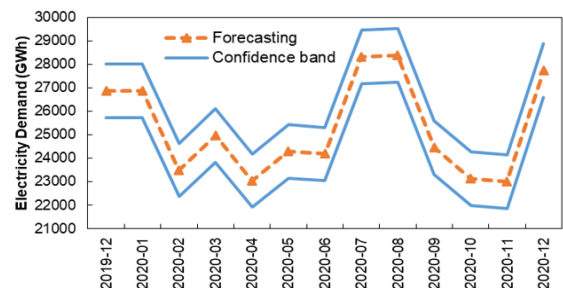


Figure 6. Forecasting results for the period of 2019/12 – 2020/12

Table 3. Comparison of the performance criteria of the studies

Performance criteria	Symbol	This study	*Hamzacebi et al. (2019)
Maximum negative error	MNE	-1746	-4872
Maximum negative percentage error	MNPE	-6.54 %	-16.48 %
Maximum positive error	MPE	1357	457
Maximum positive percentage error	MPPE	5.95 %	1.92%
Mean absolute error	MAE	596.92	2229.58
Mean absolute percentage error	MAPE	2.45 %	9.05 %
Root mean square error	RMSE	735.96	2573.13
Root mean square percentage error	RMSPE	3.00 %	10.24 %
Correlation	R	0.943	0.784
Coefficient of determination	R ²	0.877	0.615

*Performance criteria were calculated based on the forecast results given in the mentioned literature study.

The concept of GMDH-type NN might be a good tool to estimate the future electricity demand of Turkey. All energy-related institutions, especially the Turkish Ministry of Energy and Natural Resources may concern to use this research methodology in terms of strategic planning. In this way, this study can provide contribution not only to the academic research area but also to the practical life.

5. Conclusions

In this work, the GMDH-type NN approach was utilized to develop forecast model for the monthly electricity demand in Turkey. The developed forecasting model was established on the time series of the monthly electricity demand for the period of 2003/01-2018/12. The developed model was tested for the period of 2019/01-2019/11. In addition, to show the generalization capability, the forecasting results were compared with a literature study. The obtained outcomes from this work can be summarized below:

- In the model, development stage, main performance criteria such as MAPE, RMSPE and R^2 were obtained as 2.12 %, 2.67 % and 0.965, respectively.
- In the testing stage, the obtained forecasting performances were 2.10 % for MAPE, 2.36 % for RMSPE, and 0.869 for R^2 .
- In the literature comparison, the used GMDH-type NN model provided higher forecasting performance compared to the seasonal ANN model used by the literature study. While the literature study provided 9.05 % for MAPE, 10.24 % for RMSPE, and 0.784 for R^2 , GMDH-type NN model provided 2.45 %, 3 % and 0.943 for the same performance criteria, respectively.
- Based on the developed model, forecasting monthly electricity demand for the period of 2019/12 – 2020/12 were presented.

In light of the above results, it can be suggested that the GMDH type NN approach is appropriate to forecast the Turkey's monthly electricity demand. In future studies, the model forecasting capability can be enhanced by adding different input variables that influence the monthly electricity demand, such as price index, import and export data.

Declaration

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article. The author also declared that this article is original, was prepared in accordance with international publication and research ethics, and ethical committee permission or any special permission is not required.

Author Contributions

A.V. Akkaya is responsible for all section of the study.

References

1. Tutun S, Chou C-A, Caniyılmaz E, *A new forecasting framework for volatile behavior in net electricity consumption: A case study in Turkey*. Energy, 2015. 93: p. 2406-2422.
2. Cunkas M, Taskiran U, *Turkey's electricity consumption forecasting using genetic programming*. Energy Sources, Part B: Economics, Planning, and Policy, 2011. 6: p. 406–416.
3. Toksari M. D., *A hybrid algorithm of ant colony optimization (ACO) and iterated local search (ILS) for estimating electricity domestic consumption: Case of Turkey*. Electrical Power and Energy Systems, 2016. 78: p. 776–782.
4. Günay ME., *Forecasting annual gross electricity demand by artificial neural networks using predicted values of socio-economic indicator sand climatic conditions: Case of Turkey*. Energy Policy, 2016. 90: p. 92–101.
5. Hamzacebi C, Es H.A., *Forecasting the annual electricity consumption of Turkey using an optimized grey model*. Energy, 2014. 70: p. 165-171.
6. Kavaklioglu K, Ceylan H, Ozturk H.K, Canyurt O.E., *Modeling and prediction of Turkey's electricity consumption using Artificial Neural Networks*. Energy Conversion and Management, 2009. 50: p. 2719–2727.
7. Kiran MS, Ozceylan E, Gunduz M, Paksoy T. *Swarm intelligence approaches to estimate electricity energy demand in Turkey*. Knowledge-Based Systems, 2012. 36: p. 93–103.
8. İlseven E, and Göl M., *Medium-term electricity demand forecasting based on MARS*, in IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe) 2017: Torino, Sept 26-29.
9. Hong T, Fan S., *Probabilistic electric load forecasting: A tutorial review*. International Journal of Forecasting, 2016. 32: p. 914-938.
10. Yumurtaci Z, Asmaz E., *Electric energy demand of Turkey for the year 2050*. Energy Source, 2004. 26: p. 1157-1164.
11. Erdogdu E., *Electricity demand analysis using cointegration and ARIMA modelling: a case study of Turkey*. Energy Policy, 2007. 35: p. 1129-1146.
12. Dilaver Z, Hunt LC., *Turkish aggregate electricity demand: an outlook to 2020*. Energy, 2011. 36: p. 6686-6696.
13. Hamzacebi C., *Improving artificial neural networks' performance in seasonal time series forecasting*. Information Sciences, 2008. 178: p.4550–4559.
14. Saab S, Badr E, Nasr G., *Univariate modeling and forecasting of energy consumption: the case of electricity in Lebanon*. Energy, 2001. 26(1): p.1–14.
15. Bianco V, Manca O, Nardini S., *Linear regression models to forecast electricity consumption in Italy*. Energy Sour B Econ Plan Policy, 2013. 8(1): p.86–93.
16. Pao H.T. 2006. *Comparing linear and nonlinear forecasts for Taiwan's electricity consumption*. Energy 31(12):2129–2141.
17. Çunkaş M, Altun A.A., *Long term electricity demand forecasting in Turkey using artificial neural networks*. Energy Sources, Part B: Economics, Planning, and Policy, 2010. 5(3): p. 279-289.
18. Kucukali S, Baris K., *Turkeys short-term gross annual electricity demand forecast by fuzzy logic approach*. Energy Policy, 2010. 38(5): p. 2438-2445.

19. Kavaklioglu K., *Modeling and prediction of Turkey's electricity consumption using support vector regression*. Applied Energy, 2011. 88: p. 368-375.
20. Sözen A, Isikan O, Menlik T, Arcaklioglu E., *The Forecasting of Net Electricity Consumption of the Consumer Groups in Turkey*. Energy Sources, Part B: Economics, Planning, and Policy, 2011. 6(1): p. 20-46.
21. Assareh E, Behrang MA, Ghanbarzadeh A., *Forecasting energy demand in Iran using Genetic Algorithm (GA) and Particle Swarm Optimization (PSO) methods*. Energy Sources, Part B: Economics, Planning, and Policy, 2012. 7(4): p.411-422.
22. Ayvaz B, Kusakci A.O., *Electricity consumption forecasting for Turkey with nonhomogeneous discrete grey model*. Energy Sources, Part B: Economics, Planning, and Policy, 2017. 12(3): p. 260-267.
23. Bilgili M, Sahin B, Yasar A, Simsek E., *Electric energy demands of Turkey in residential and industrial sectors*. Renewable and Sustainable Energy Reviews, 2012. 16(1): p. 404-414.
24. Chen JF, Lo SK, Do QH., *Forecasting monthly electricity demands: an application of neural networks trained by heuristic algorithms*. Information, 2017. 8(1): 31.
25. Uzlu E., *Application of Jaya algorithm-trained artificial neural networks for prediction of energy use in the nation of Turkey*. Energy Sources, Part B: Economics, Planning, and Policy, 2019. 14(5): p. 183-200.
26. Hamzacebi C, Es HA, Cakmak R., *Forecasting of Turkey's monthly electricity demand by seasonal artificial neural network*. Neural Computing and Applications, 2019. 31: p. 2217-2231.
27. Luzar M, Sobolewski Ł, Miczulski W., Korbicz J., *Prediction of corrections for the Polish time scale UTC(PL) using artificial neural networks*. Bulletin of the Polish Academy of Sciences Technical Sciences, 2013. 61(3): p. 589-594.
28. Amjady N, Daraeepour A., *Midterm Demand Prediction of Electrical Power Systems Using a New Hybrid Forecast Technique*. IEEE Transactions on Power Systems, 2011. 26(2): p. 755-765.
29. Ivakhnenko AG., *Polynomial theory of complex systems*, IEEE Transactions on Systems, Man and Cybernetics, 1971. 1(4): p. 364-378.
30. Ahmadi MA, Golshadi M., *Neural network based swarm concept for prediction asphaltene precipitation due to natural depletion*. Journal of Petroleum Science and Engineering, 2012. 98: p. 40-49.
31. Ahmadi MH., Ahmadi MA, Mehrpooya M, Rosen MA., *Using GMDH neural networks to model the power and torque of a Stirling engine*. Sustainability, 2015. 7 (2): p. 2243-2255.
32. Nariman-Zadeh, N, Darvizeh A, Felezi ME, and Gharababaei H., *Polynomial modelling of explosive compaction process of metallic powders using GMDH-type neural networks and singular value decomposition*. Modelling and Simulation in Materials Science and Engineering, 2002. 10 (6): 727.
33. Rezaei MH, Sadeghzadeh M, Nazari MA, Ahmadi MH, Astarai FR., *Applying GMDH artificial neural network in modeling CO2 emissions in four Nordic countries*, International Journal of Low-Carbon Technologies, 2018. 13: p. 266-271.
34. Jia X, Di Y, Feng J, Yang Q, Dai H, Lee J., *Adaptive virtual metrology for semiconductor chemical mechanical planarization process using GMDH-type polynomial neural networks*. Journal of Process Control, 2018. 62: p. 44-54.
35. Zhou L., *Prediction of CO2 adsorption on different activated carbons by hybrid group method of data-handling networks and LSSVM*, Energy Sources, Part A: Recovery, Utilization, and Environmental Effects, 2019. 41 (16): p. 1960-1971.
36. Li RYM, Fong S, Chong KWS., *Forecasting the REITs and stock indices: Group Method of Data Handling Neural Network Approach*. Pacific Rim Property Research Journal, 2017. 23(2): p. 123-160.
37. Turkish Electricity Transmission Cooperation (TEIAS). 2019. Electricity Generation Transmission Statistics of Turkey. [cited 2019, 21 December]; Available from: <http://www.teias.gov.tr>.



e-ISSN: 2618-575X

INTERNATIONAL ADVANCED RESEARCHES
and
ENGINEERING JOURNAL

Journal homepage: www.dergipark.org.tr/en/pub/iarejInternational
Open Access Volume 05
Issue 01

April, 2021

Research Article**Mechanical characterization of pack-boronized AISI 4140 and AISI H13 steels****Dilek Arslan** ^{a,*}  and **Selda Akgün Kayral** ^b ^aGraduate School of Applied and Natural Sciences, Department of Mechanical Engineering, Manisa Celal Bayar University, 45140, Manisa, Turkey^bHasan Ferdi Turgutlu Faculty of Technology, Department of Mechanical Engineering, Manisa Celal Bayar University, 45400, Manisa, Turkey

ARTICLE INFO

Article history:

Received 27 October 2020

Revised 20 December 2020

Accepted 05 January 2021

Keywords:

Microhardness

Pack boronizing

Surface roughness

Wear resistance

ABSTRACT

Wear losses have a great importance in the world machinery industry. They cause billions of dollars in financial losses every year. Studies on surface treatments are increasing day by day in order to minimize the wear losses of materials. In this study, the pack boronizing process was applied to AISI 4140 and AISI H13 steels, which are frequently used in the manufacturing and molding industry, by using Ekabor II powder at 900 °C and 950 °C for 4 and 6 hours. Microstructural examinations of the samples subjected to metallographic processes were carried out. Afterwards, microhardness measurements were performed by applying 50 gf load for 10 seconds. Wear tests were carried out using pin-on-disk tribotests in a dry environment under 2 N and 5 N loads on the CSM Tribometer device. Wear losses were measured as volumetric loss. Thanks to the boronizing process, surface quality, surface hardness, and wear resistance of both steel materials were increased at a high rate.

© 2021, Advanced Researches and Engineering Journal (IAREJ) and the Author(s).

1. Introduction

In today's industry, significant investments are made in steel materials which have a wide application area. However, these materials have short service lives due to various wear caused by both mechanical factors and oxidative and corrosive ambient conditions. Material losses caused by wear lead to billions of dollars in financial losses each year in the world machinery industry. Therefore, the problem of wear should be taken into account not only in its technical aspect but also in its negative impact on the economy. In order to prevent these losses and to extend the service life of the steels, an appropriate surface coating process should be applied to these steels in addition to the correct design and appropriate steel selection. The surface coating has layers with high hardness and abrasion resistance by pack-boronizing on metallic surfaces. This thermochemical surface treatment is also known as a surface hardening method, and it has a wide usage area in the industry due to its easy applicability and economy. Boride layers with

high mechanical performance are formed on the surface of steels thanks to boronizing based on the diffusion of boron atoms to the base material surface. It has been reported in various studies that the friction coefficient of the borided surfaces obtained after the boriding process based on the thermal diffusion of boron atoms is very low [1-3]. In addition, thanks to the surface hardness that reaches up to 2000 HV, which it gives to the surface of the material, boronizing is seen as a solution against abrasive or adhesive wear [4-10]. Moreover, even at temperatures well above room temperature (up to 650 °C), the material surface can maintain this high hardness [8]. Boronized steels have high hardness (about 2000 HV) and high wear resistance [7, 11]. The combination of high surface hardness and low friction coefficient of the boron layer obtained on the substrate material surface by boronizing contributes significantly to the fight against wear [12]. AISI 4140 tempered steel is a medium carbon and low alloy steel material, and thanks to its high wear resistance and high toughness and a good hardness-ductility balance, it has a very wide usage area in automotive (e.g.,

* Corresponding author. Tel.: +90(236) 201 27 12; Fax: +90(236) 201 27 15

E-mail addresses: da.dilekarslan@gmail.com (D. Arslan), selda.akgun@cbu.edu.tr (S. Akgün Kayral)

ORCID: 0000-0003-0198-0787 (D. Arslan), 0000-0003-1971-1550 (S. Akgün Kayral)

DOI: 10.35860/iarej.817274

This article is licensed under the CC BY-NC 4.0 International License (<https://creativecommons.org/licenses/by-nc/4.0/>).

crankshaft, axle shaft, and gear cone, machining processes; machine tools, parts such as rams, spindles, studs, bolts, and nuts) [13, 14]. AISI H13 hot work tool steel, which is frequently used in the diesinking industry, is a high-strength and ductility tool material used especially in extrusion and injection molds of metals, plastic injection molds, hot pressing of copper alloys and steel forging dies [15, 16].

Both steel materials are exposed to high temperature and pressure due to the working environments in which they are used. Therefore, not only the service life will be prolonged by preventing deformation and wear on the surface of these materials, but also the expected high performance from the materials in aggressive working conditions will be achieved. Some studies have focused on improving the wear resistance of hot work tool steels to make them cope with these challenging conditions that they are exposed to [16-20]. The developments in industrial applications in recent years have led to the need for materials with superior properties [21]. Boride layer formed on matrix material surface with boronizing shall also serve as a thermal barrier on the surface of the machine elements operating under high temperature. Thus, direct contact of the material surface with mechanical and tribological factors will be prevented. In this way, boronizing will be an economical alternative method to thermal barrier coating (TBC), which requires a very complex assembly and high-cost initial investment. It has been reported that this thermal barrier coating formed on the surface of the machine elements reduces wear and significantly improves the performance of the machine element and its mechanism [22, 23]. In many studies in the literature, the effect of boronizing on hardness and wear in AISI 4140 and AISI H13 steels has been examined. Ulutan et al. [24] applied the pack boronizing process to AISI 4140 steel by using commercial EKabor2[®] powder for 2, 4, and 6 hours at 900, 950, 1000, and 1050 °C. The results of the study showed that compared to untreated samples, the hardness of boronized AISI 4140 steel increased approximately 6 times, while the abrasive wear resistance increased approximately 3-4 times, and the friction coefficient was very low. Kara et al. [15] came to the conclusion that the hardness of the material increased from 485 HV_{0.05} to 1989 HV_{0.05} as a result of the pack boronizing process applied to AISI H13 steel at 900 and 950 °C for 2, 4, and 6 hours. Cimenoglu et al. [25] investigated the tribological behavior of the boron layer obtained on AISI 4140 steel by pack boronizing at high temperature. As a result of the pack boronizing process carried out using EKabor2[®] powder for 12 hours at 750, 800, 850, and 900 °C temperatures, they obtained a very good wear resistance in all samples at room temperature; this superior wear resistance obtained in samples boronized at 850 °C and 900 °C and consisting of a double phase boron layer

(Fe₂B+FeB) was also maintained at 300 °C. Cárdenas et al. [26] applied pack boronizing to AISI H13 tool steel at 1000 °C for 8 hours and examined the wear behavior by using ball-on-disc method. With boronizing, the hardness of AISI H13 steel reached 1803 HV and the friction coefficient decreased from 0.3 to 0.15. While cracks and grinding were observed on the unboronized sample surface after the wear test, only non-significant shear wear marks were observed on the boronized sample surfaces. The shear wear resistance of AISI H13 steel increased 13 times after boronizing.

In this study, pack boronizing was performed on AISI 4140 and AISI H13 steels at 900 °C and 950 °C for 4 and 6 hours. After the boronizing process, the microstructural analyses of the samples were performed and surface roughness, microhardness, and wear strengths were measured. With increased strength and wear resistance, steel materials whose surface and mechanical properties are improved with pack boronizing will provide significant advantages to mold manufacturers and end-users. The most important of these advantages will be the economic gain to be achieved due to the prolongation of the mold and tool life of steel materials. In addition, as an alternative to hot work tool steels used in high-temperature applications, surface quality, hardness, and wear resistance were increased by boronizing, and a boronized AISI 4140 tempered steel with a lower cost was also proposed.

2. Experimental Details

2.1 Pack Boronizing Process

The chemical compositions of AISI 4140 tempered steel and AISI H13 hot work tool steel used in the study are shown in Table 1. Samples to be boronized were prepared in Ø12x7mm size shaped as discs. Before the boronizing process, the samples were polished, cleaned with ultrasonic bath, and air-dried, then placed in steel boxes filled with Ekabor II powder in a way that they would not contact each other, and covered with boronizing powder again. Before the lid was closed, it was completely filled with SiC powder to prevent boronizing powder from being affected by the air in the oven. Afterwards, the lid of the boxes was closed and the surroundings were plastered with mud to prevent air from getting into them. The pack boronizing process was carried out in a laboratory type ash oven at 900 °C and 950 °C for 4 and 6 hours. After the boronizing process was completed, the steel crucibles taken out from the oven were cooled at room temperature and the samples were removed from the crucible.

Table 1. Chemical composition of the samples (wt.%)

	C	Si	Mn	Cr	Mo	V
AISI H13	0.40	1.00	0.35	5.15	1.40	1.00
AISI 4140	0.38-0.45	0.15-0.40	0.50-0.80	0.90-1.20	0.15-0.30	-

2.2 Surface Roughness Measurement

Measurements of the surface roughness were performed with Mitutoyo SJ 301 brand roughness measuring device before and after the boronizing process to determine the effect of the boronizing process on surface roughness. This device was also used to extract wear surface profiles during the volumetric measurement of wear losses.

2.3 Characterization of Boride Layers

Before the microstructural analyses, the samples were sanded with 240, 400, 600, 800, 1000, and 1200 grit mesh sanding sheets respectively after molding in the hot bakelite extraction device for ease of grip. It was then polished with 3 μm Struers DiaPro MoI₃B and 1 μm Struers DiaPro NapB1 solutions. Polished samples were examined with the NIKON Eclipse LV100 brand optical microscope. In the examinations of the coating layers, the coating layer thicknesses were measured using the Clemex brand image analysis program of the optical microscope.

2.4 Surface Roughness Measurement

Microhardness measurements of the samples with polished surfaces were performed by applying 50 gf load for 10 seconds on Future-Tech FM-700 brand device. Hardness measurements were made along a single line from the beginning of the coating to the base metal. Thus, in addition to the coating thickness, the hardness depth was also measured. The optimum parameter was obtained by keeping the samples at 950 °C for 6 hours after the microhardness process.

2.5 Wear Tests

In the study, the dry sliding wear tests were performed on the test samples which were pack boronized at 950 °C for 6 hours and unboronized by using CSM Tribometer Device. The dry sliding wear tests were performed in pin-on-disc assembly under dry friction conditions. WC balls with a diameter of 6 mm were used as the counter element. Wear tests were carried out at dry sliding paths of 250 m and 500 m at velocity of 11.77 cm/sec by applying 2 N and 5 N loads. Abrasion losses were calculated as volumetric loss. The profile of the wearing surface was measured with the surface roughness measuring device and volumetric losses were calculated.

3. Results and Discussion

3.1 Surface Roughness

The surface roughness of the samples was measured before the boronizing process and after 6 hours of boronizing at 950 °C, and the results shown in Table 2 were obtained. When the measurement results in Table 2 are examined, it is seen that the boronizing process has a positive effect on surface roughness. Besides the boronizing process is known to increase surface roughness values in the boronizing of

Table 2. Effect of Boronizing on Surface Roughness

	4140	B-4140	H13	B-H13
Ra (μm)	1.88	1.59	1.96	1.34

precision machined surfaces, it also causes the boronizing surface quality to increase in materials with higher surface roughness. The results confirmed this situation. Therefore, the use of boronizing in manufacturing will shorten the time of the machining process and thus reduce the production costs.

3.2 Characterization of Boride Layers

Microstructural images were taken after etching the polished samples with 4% nital solution. Figure 1(a) shows the microstructural images of AISI 4140 steel boronized at 900 °C for 4 hours. When the boronizing layers are examined, a structure with porous but having a columnar morphology is encountered as in the literature [27]. Since the outermost surface of the coating is damaged due to the polishing process, it seems to be porous. Due to the alloy elements in the chemical component of AISI 4140 material, the boronizing layer has not been able to develop further. Figure 2(a) shows the microstructural images of AISI H13 steel boronized at 900 °C for 4 hours. Since the alloy ratio of AISI H13 material is higher than AISI 4140 material, it is seen that the coating layer development is weaker at the same temperature and time. The growth of the coating did not occur columnar due to the density of the alloy elements in agreement with the literature [28].

When Figure 1(b) is examined, it can be concluded that the boride layer of AISI 4140 steel boronized at 900 °C for 6 hours is porous and continuous, and that FeB phase formation occurs due to the increase in diffusion time outside Fe₂B phase compared to Figure 1(a). When the microstructural images of AISI H13 steel boronized at 900 °C for 6 hours in Figure 2(b) are compared with Figure 2(a), it is observed that the increased diffusion time causes an increase in the development of the coating, and a nonporous layer formation occurs compared to the 4-hour coating. In addition, FeB phase formation other than Fe₂B phase occurred with the increase in time. In Figure 1(c), the boride layer of AISI 4140 steel boronized at 950 °C for 4 hours is seen as a partially columnar but porous boride layer, and no significant effect of boronizing temperature on the coating layer has been observed compared to Figure 1(a). Examining Figure 2(c), it is understood that after boronizing process performed in AISI H13 steel at 950 °C for 4 hours, the coating layer does not develop in columnar morphology and the boride layer has a porous structure. When compared with Figure 2(a), it is understood that boronizing temperature does not have a positive effect on boronizing behavior. As seen in Figure 1(d), although the boride layer of AISI 4140 steel boronized at 950 °C for 6 hours developed in the columnar morphology, it was extremely porous. Due to the increase in temperature and duration, a homogeneous

thickness and columnar-developed FeB phase were also encountered. The fact that the FeB phase was homogeneously distributed and had equal thickness caused the wear resistance of the material to increase. If the formation of the FeB phase were to occur in irregular dimensions, this would certainly be an undesirable problem causing a rapid increase in wear losses. However, in this parameter, the columnar development of the layer created a multi-layer coating effect. In Figure 2(d), it is observed that the AISI H13 steel boronized at 950 °C for 6 hours has developed in a homogeneous and porous structure. As a result of the boronizing process, no FeB phase was found.

Boride layer thickness measurement was measured from 5 different points in Clemex brand image analysis program and the results presented in Table 3 were obtained by taking the average values. When Table 3 is examined, it can be seen that the coating layer values obtained with the comments made in the microstructural images confirm each other. It is seen that the increase in the ratio of alloy element in the material has a negative effect on the development of the boride layer. For both materials, the highest boride layer thicknesses were reached by boronizing performed at 950 °C for 6 hours.

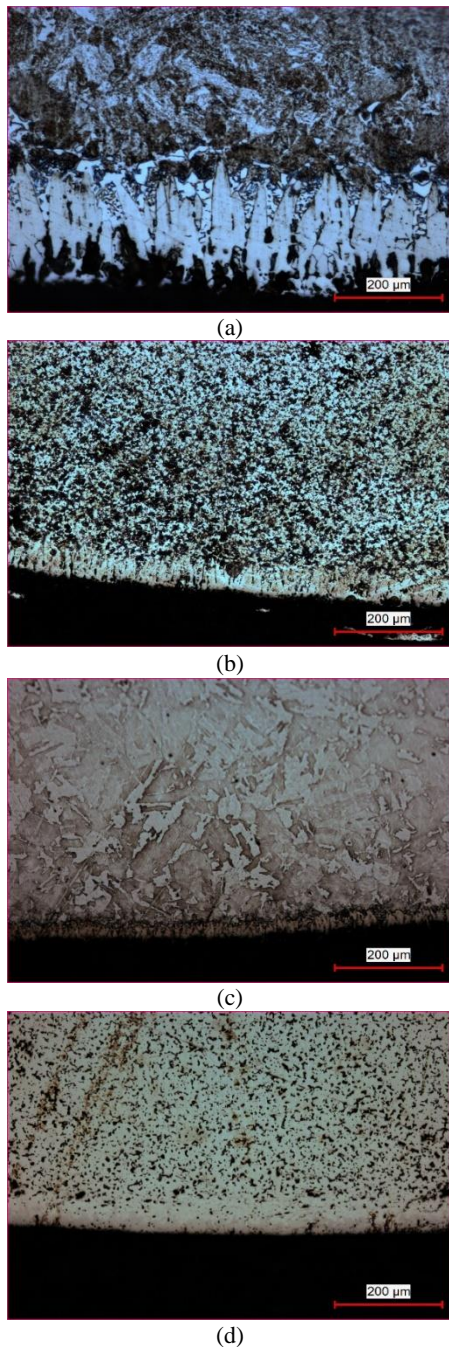


Figure 1. Microstructural images of AISI 4140 boronized a) at 900 °C for 4 hours b) at 900 °C for 6 hours c) at 950 °C for 4 hours and d) at 950 °C for 6 hours

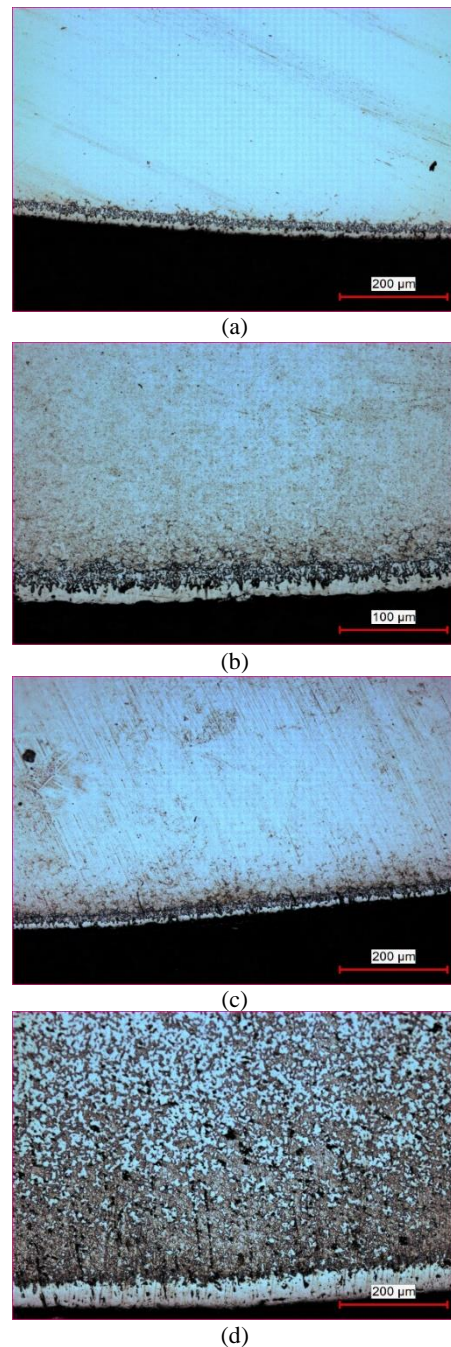


Figure 2. Microstructural images of AISI H13 boronized a) at 900 °C for 4 hours b) at 900 °C for 6 hours c) at 950 °C for 4 hours and d) at 950 °C for 6 hours

Table 3. Boride Layer Measurement Results

	Thickness (μm)
B-4140-900 °C-4h	40
B-H13-900 °C-4h	35
B-4140-950 °C-4h	60
B-H13-950 °C-4h	35
B-4140-900 °C-6h	40
B-H13-900 °C-6h	40
B-4140-950 °C-6h	80
B-H13-950 °C-6h	60

3.3 Microhardness

Microhardness measurements were carried out in such a way that they were closest to the starting point of the coating and the distance between the traces was at least 4 times the trace size. Figure 3 shows the microhardness changes of the boride layer of AISI 4140 material boronized at different temperatures and times. As the boronizing temperature and time increase, both the hardness and hardness depth of the AISI 4140 material (i.e. the diffusion layer) increase. The evaluation made with the microstructural images indicated that the highest hardness values (over 1200 HV) were reached in the parameters in which the FeB phase, which corresponded to the hardness values of the FeB phase, occurred. It is seen that the highest hardness and hardness depth has been reached in samples boronized at 950 °C for 6 hours. Table 3 shows that this corresponds to the highest boride layer.

Figure 4 shows the microhardness change of the boride layer of AISI H13 steel boronized at different temperatures and times. The examination of microstructural images of the boronized AISI H13 steel indicates that the FeB phase

definitely has not occurred. The fact that the highest microhardness value can reach up to 1100 HV is due to the fact that the FeB phase, which significantly increases the hardness in the boride layer, does not occur. Examining Table 3, it is seen that the thickness of the coating layer is in parallel with Figure 4. When Figure 3 and Figure 4 are examined together, it is figured out that the diffusion zone is not deep after passing the boride layer. The main reason for this is the alloy contents of both materials.

3.4 Abrasion Resistance

Figure 5 shows the graph of abrasion loss by material under 2N load. It is seen that the unboronized AISI H13 material has suffered from the greatest abrasion loss under 2N load. Abrasion losses of the boronized materials have reduced significantly. Abrasion loss after the boronizing process of the AISI 4140 material has decreased by 75%. It is seen that the decrease in abrasion loss reaches 83% in AISI H13 material. The best wear resistance has been achieved in boronized AISI 4140 material. The examination of microstructural images indicates that the FeB and Fe₂B phases have caused this. The fact that the abrasion losses do not increase despite the formation of the FeB layer is proof that the FeB layer is connected to the substrate with a good adhesion force.

Figure 6 indicates the abrasion losses of the materials due to the sliding path. AISI H13 suffered the highest abrasion loss at a sliding path of 250m under a 2N load, while boronized materials showed the same abrasion losses. At a sliding path of 500m, boronized AISI 4140 material showed the best wear resistance due to its high surface hardness and the depth of the diffusion layer.

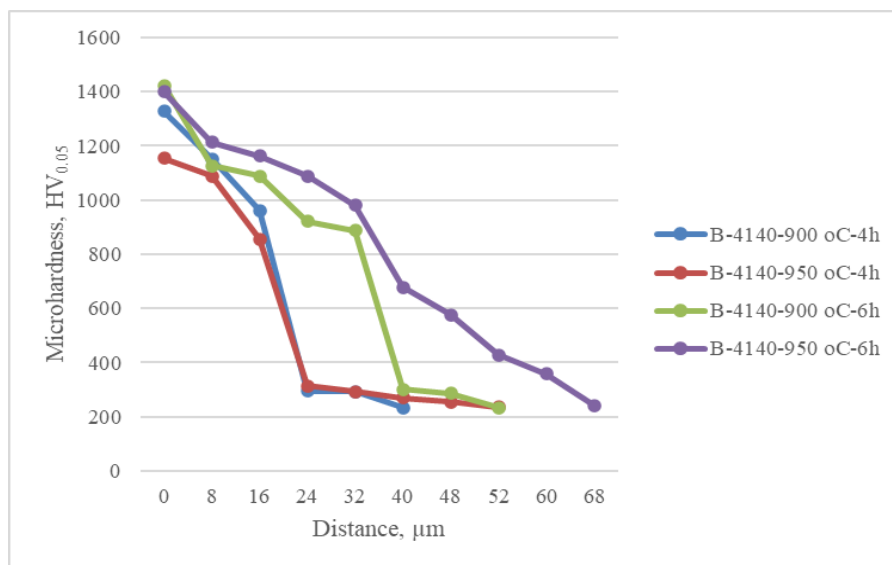


Figure 3. The change in microhardness as a function of depth below the surface of AISI 4140 steel samples boronized at different temperatures and times

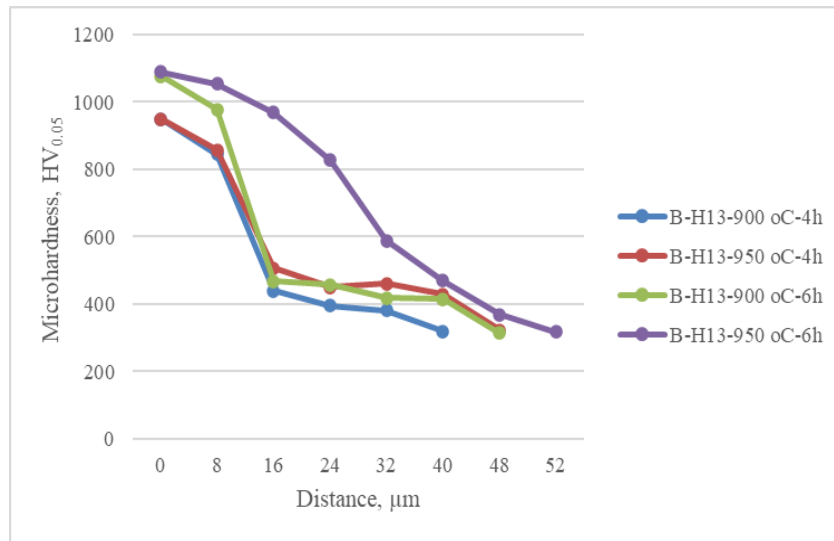


Figure 4. The change in microhardness as a function of depth below the surface of AISI H13 steel samples boronized at different temperatures and times

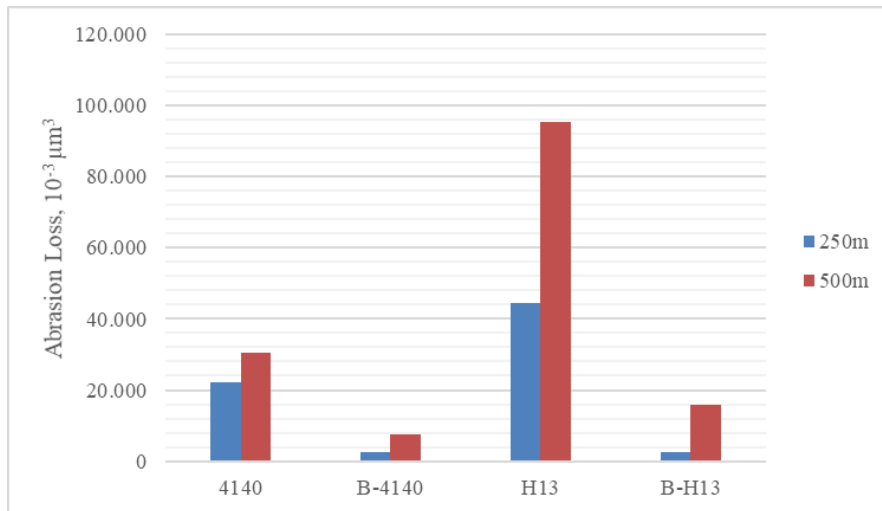


Figure 5. Abrasion losses of samples by material (under 2N load)

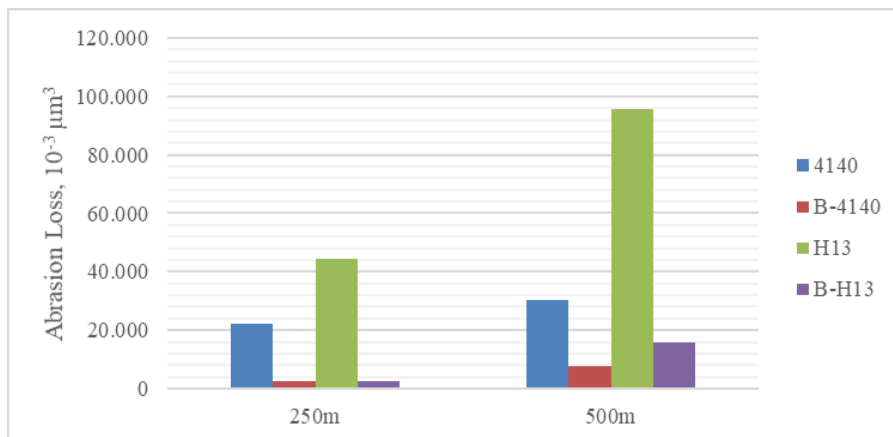


Figure 6. Abrasion losses of samples by sliding path (under 2N load)

Figure 7 shows the abrasion loss graph according to the material under 5N load. It is seen that AISI H13 material,

which has not been boronized, has suffered from the greatest abrasion loss under 5N load. The abrasion loss of

it is nearly 2 times the abrasion loss of the unboronized AISI 4140 material. Abrasion losses of boronized materials have decreased significantly. Abrasion loss of AISI 4140 material decreased by 75% after boronizing. It is seen that the decrease in abrasion loss has reached 73% for AISI H13 material. The best wear value has been achieved in boronized AISI 4140 material. When the microstructural images are examined, it is understood that the FeB and Fe₂B phases have caused this. The fact that the abrasion losses do not increase despite the formation of the FeB layer is proof that the FeB layer is connected to the substrate with a good adhesion force. Figure 8 shows the materials' adhesion losses depending on the sliding path under 5N load. The unboronized AISI H13 was subjected to the highest abrasion loss at a sliding path of 250m, while the boronized AISI 4140 showed the least abrasion loss.

At a sliding path of 500m, boronized AISI 4140 material showed the best wear resistance due to its high surface hardness and the depth of the diffusion layer. In Figure 9-22, the wearing surfaces of the materials and the wearing surfaces of the counter elements are displayed with images taken by the Stereo Microscope.

These images indicate that the abrasion losses are high on the unboronized samples and adhesions from the unboronized material to the counter element occur. The width of the wear trace is also increasing. It is clearly seen in Figure 13 that there is an adhesive wear especially on the wear of unboronized AISI H13 material. Adhesive wear occurred especially on all unboronized samples. Because it is well known that the boride layer plays a significant role in improving the resistance to abrasion and adhesion [29, 30].

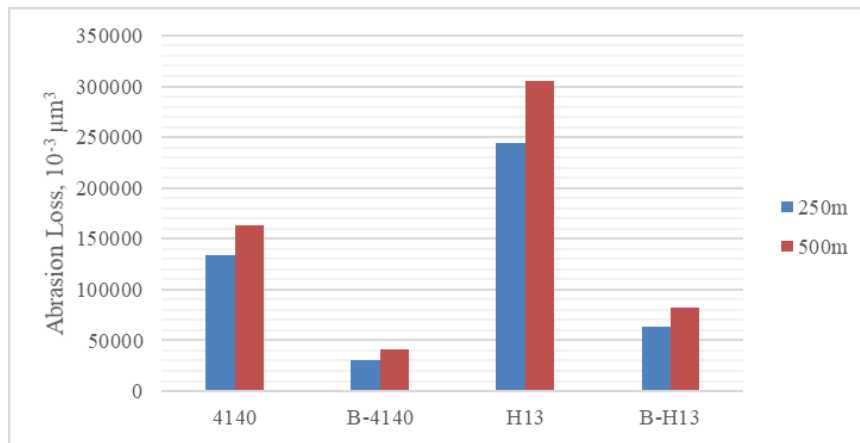


Figure 7. Abrasion loss of samples by material (under 5N load)

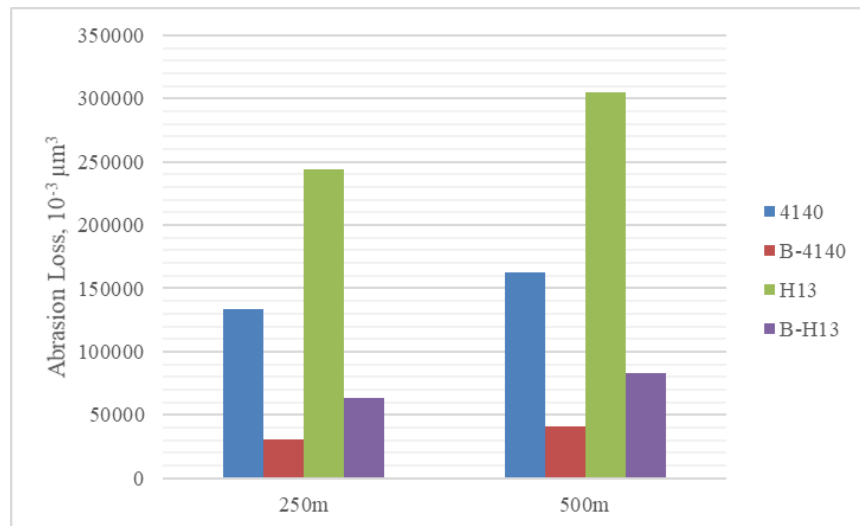


Figure 8. Abrasion loss of samples by sliding path (under 5N load)

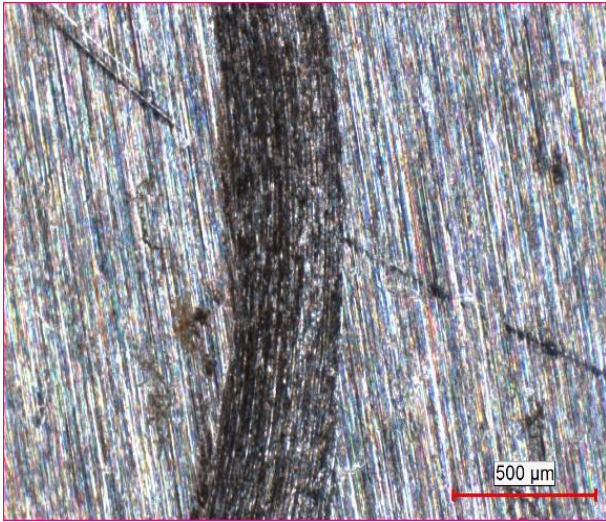


Figure 9. Wearing surface of AISI 4140 material abraded under 2N load

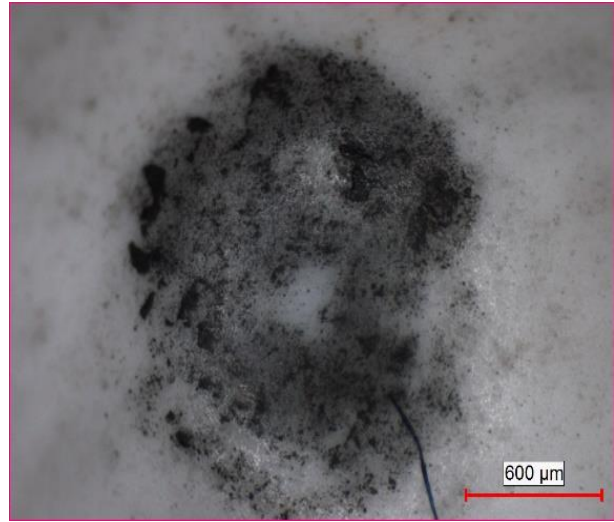


Figure 12. Wearing surface of counter element of boronized AISI 4140 material abraded under 2N load



Figure 10. Wearing surface of counter element of AISI 4140 material abraded under 2N load

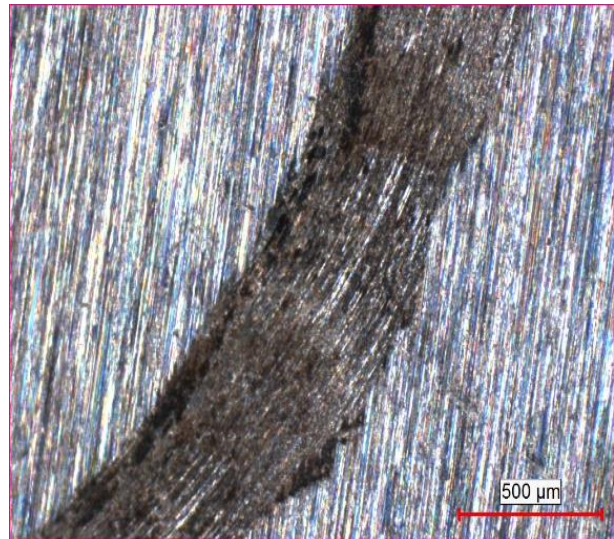


Figure 13. Wearing surface of AISI H13 material abraded under 2N load

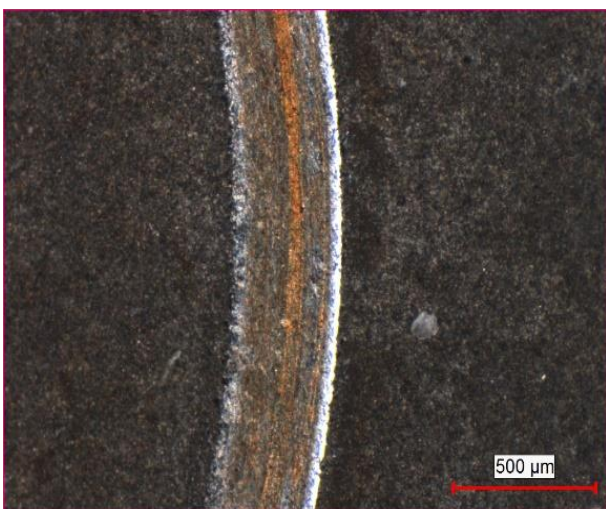


Figure 11. Wearing surface of boronized AISI 4140 material abraded under 2N load



Figure 14. Wearing surface of counter element of AISI H13 material abraded under 2N load

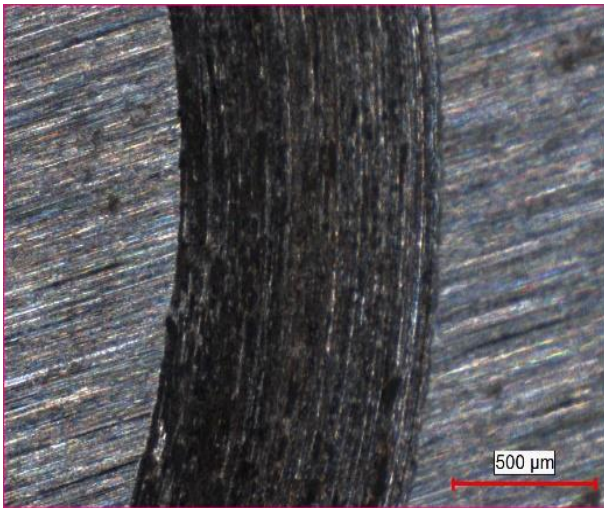


Figure 15. Wearing surface of AISI 4140 material abraded under 5N load

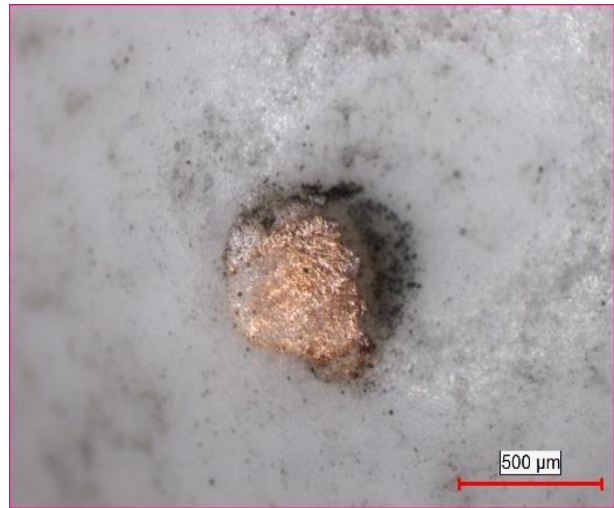


Figure 18. Wearing surface of counter element of boronized AISI 4140 material abraded under 5N load



Figure 16. Wearing surface of counter element of AISI 4140 material abraded under 5N load

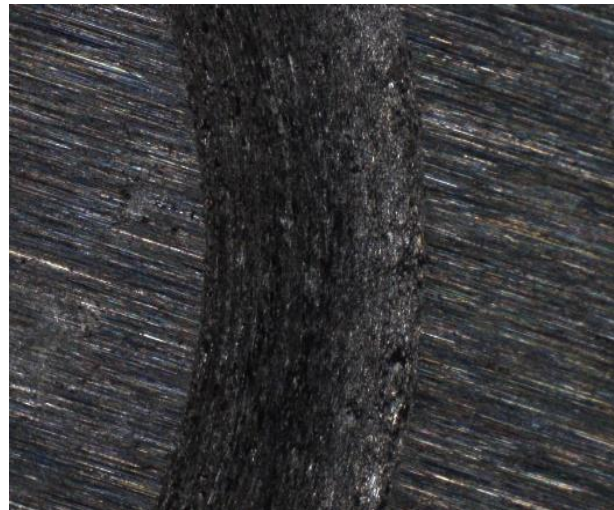


Figure 19. Wearing surface of AISI H13 material abraded under 5N load

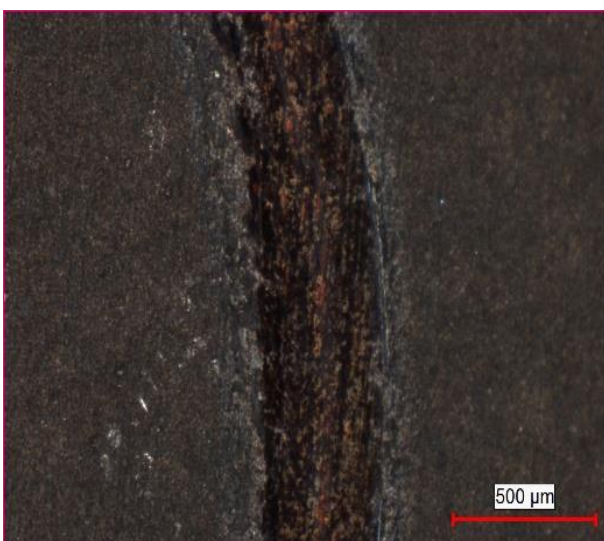


Figure 17. Wearing surface of boronized AISI 4140 material abraded under 5N load

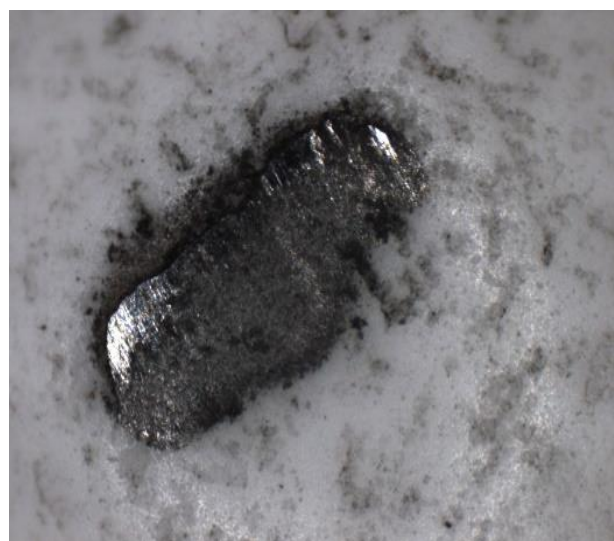


Figure 20. Wearing surface of counter element of AISI H13 material abraded under 5N load

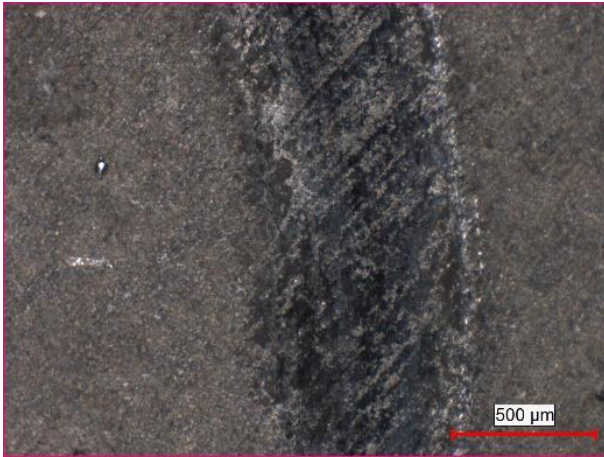


Figure 21. Wearing surface of boronized AISI H13 material abraded under 5N load



Figure 22. Wearing surface of counter element of boronized AISI H13 material abraded under 5N load

4. Conclusions

Within the scope of this study, boronizing process was performed on AISI 4140 and AISI H13 steel materials and appropriate boronizing parameter was determined. In order to determine the appropriate boronizing parameter, the materials were subjected to boronizing process at different times and temperatures. Optimum boronizing parameters were reached as a result of microstructural examinations and microhardness tests. Common parameters for both types of materials are determined as 950 °C temperature and 6 hours retention time. Surface roughness, hardness, and wear tests were applied to the samples, which were boronized with these parameters by using the pack boronizing process. The results and recommendations are given below.

- FeB phase was formed on AISI 4140 material boronized for 6 hours at both temperature values. This layer is highly homogeneous and columnar dependent on the substrate material.
- Boride layer thickness of 80 μm was reached in AISI 4140 material and 60 μm in AISI H13 material boronized at 950 °C for 6 hours.
- Microhardness values up to 1400 HV were reached in the process parameters where FeB phase occurred.

- As a result of the wear tests, the material with the lowest wear resistance was determined as AISI H13. However, by applying boronizing process, the properties of this material can be improved by 75%.
- The material with the highest wear resistance was determined as boronized AISI 4140. Thanks to the duplex coating on the surface, the wear resistance was improved up to 70%.
- The comparison of the wear resistance of the two boronized materials indicated that the boronized AISI 4140 material showed 2 times more strength than the boronized AISI H13.

In this study, as an alternative to AISI H13 hot work tool steel, surface quality, hardness, and wear resistance of the samples were increased by using boronizing. Finally, the choice of boronized AISI 4140 tempered steel, which is more cost-effective, is recommended.

Declaration

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article. The authors also declared that this article is original, was prepared in accordance with international publication and research ethics, and ethical committee permission or any special permission is not required.

Author Contributions

All authors conceived and conducted the study. D. Arslan. performed the experiments, analyzed the data, and wrote the paper. S. Akgün Kayral supervised and improved the study. All authors contributed to manuscript revisions. All authors read and approved the final version of the manuscript.

References

1. Budinski, K.G., *The Wear of Diffusion Treated Surfaces*, Wear, 1993. **162-164**: p. 757-762.
2. Venkataraman, B. and Sundararajan, G., *The High Speed Sliding Wear Behavior of Boronized Medium Carbon Steel*, Surface and Coatings Technology, 1995. **73**: p. 177-184.
3. Küper, A., Qiao, X., Stock, H.R. and Mayr, P., *A Novel Approach to Gas Boronizing*, Surface and Coatings Technology, 2000. **130**(1): p. 87-94.
4. Tabur, M., Izciler, M., Gul, F. and Karacan, I. *Abrasive wear behavior of boronized AISI 8620 steel*, Wear, 2009. **266**: p. 1106–1112.
5. Ulutan, M., Celik, O., Gasan, H. and Er, U. *Effect of Different Surface Treatment Methods on the Friction and Wear Behavior of AISI 4140 Steel*, Journal of Materials Science & Technology. 2010. **26**: p. 251–257.
6. Gunes, I., *Wear Behavior of Plasma Paste Boronized of AISI 8620 Steel with Borax and B₂O₃ Paste Mixtures*, Journal of Materials Science & Technology, 2013. **29**: p. 662–668.
7. Sinha, A.K., *Boriding (Boronizing)*. 1991, Materials Park USA: ASM International Handbook, The Materials

- International Society. **4**: p. 437–447.
8. Stewart, K., *Boronizing protects metals against wear*, Advanced Materials and Processes, 1997. **151**(3): p. 23–25.
 9. Gök, M., Küçük, Y., Erdoğan, A., Öge, M., Kanca, E. and Günen, A., *Dry sliding wear behavior of borided hot-work tool steel at elevated temperatures*, Surface and Coatings Technology, 2017. **328**: p. 54–62.
 10. Soydan, Y., Köksal, S., Demirer, A. and Çelik, V., *Sliding Friction and Wear Behavior of Pack-Boronized AISI 1050, 4140, and 8620 Steels*, Tribology Transactions, 2008. **51**(1): p. 74–81.
 11. Hu, R., Fenske, G.R., Rehn, L.E., Baldo, P.M., Erdemir, A., Lee, R.H. and Erck, R.A., *Tribological performance of ion-beam-mixed Fe/B multilayers on M50 steel*, Surface and Coatings Technology, 1990. **42**(3): p. 283–297.
 12. Petrova, R.S., Suwattananont, N. and Samardzic, V., *The effect of boronizing on metallic alloys for automotive applications*, Journal of Materials Engineering and Performance, 2008. **17**: p. 340–345.
 13. Lin, Y.C., Chen, M.S. and Zhong, J., *Effects of deformation temperatures on stress/strain distribution and microstructural evolution of deformed 42CrMo steel*, Materials&Design, 2009. **30**: p. 908–913.
 14. Lin, Y.C., Chen, M.S. and Zhong, J., *Effect of temperature and strain rate on the compressive deformation behavior of 42CrMo steel*, Journal of Materials Processing Technology, 2008. **205**: p. 308–315.
 15. Kara, R., Çolak, F. and Kayali, Y., *Effect of temperature and strain rate on the compressive deformation behavior of 42CrMo steel*, Transactions of the Indian Institute of Metals, 2016. **69**(6): p. 1169–1177.
 16. Taktak, S., *Some mechanical properties of borided AISI H13 and 304 steels*, Materials&Design, 2007. **28**(6), p. 1836-1843.
 17. Hernandez, S., Hardell, J., Winkelmann, H., Ripoll, M.R. and Prakash, B., *Influence of temperature on abrasive wear of boron steel and hot forming tool steels*, Wear, 2015. **338-339**, p. 27–35.
 18. Birol, Y., *Sliding wear of CrN, AlCrN and AlTiN coated AISI H13 hot work tool steels in aluminium extrusion*, Tribology International, 2013. **57**, p. 101–106.
 19. We, M., Wang, S., Wang, L. and Chen, K., *Effect of microstructures on elevated-temperature wear resistance of a hot working die steel*, Journal of Iron and Steel Research, 2011. **18**(10), p. 47-53.
 20. Birol, Y. and Isler, D., *Abrasive wear performance of AlCrN-coated hot work tool steel at elevated temperatures under three-body regime*, Wear, 2011. **270**(3-4), p. 281-286.
 21. Degirmenci, D., Kisa, M., Ozen, M. and Demircan, G., *Environmental effects on tribological behaviour of composite materials*, International Advanced Researches and Engineering Journal, 2018. **2**(3): p. 234-239.
 22. Vural, E. and Özel, S., *The investigation of effect of the ceramic coatings with bond-layer coated on piston and valve surface on engine performance of a diesel engine*, International Advanced Researches and Engineering Journal, 2020. **4**(2): p. 87-93.
 23. Özel, S., Vural, E. and Binici, M., *Taguchi method for investigation of the effect of TBC coatings on NiCr bond-coated diesel engine on exhaust gas emissions*, International Advanced Researches and Engineering Journal, 2020. **4**(1): p. 14-20.
 24. Ulutan, M., Yildirim, M.M., Çelik, O.N. and Buytoz, S., *Tribological properties of borided AISI 4140 steel with the powder pack-boriding method*, Tribology Letters, 2010. **38**: p. 231–239.
 25. Cimenoglu, H., Atar, E. and Motallebzadeh, A., *High temperature tribological behaviour of borided surfaces based on the phase structure of the boride layer*, Wear, 2014. **309**(1-2): p. 152–158.
 26. Cárdenas, E.E.V., Lewis, R., Pérez, A.I.M., Ponce, J.L.B., Pinal F.J.P., Domínguez, M.O. and Arreola, E.D.R., *Characterization and wear performance of boride phases over tool steel substrates*, Advances in Mechanical Engineering, 2016. **8**(2): p. 1–10.
 27. Keddám, M., Ortiz-Domínguez, M., Gómez-Vargas, O.A., Arenas-Flores, A., Flores-Rentería, M.A., Elias-Espinosa, M. and García-Barrientos, A., *Kinetic study and characterization of borided AISI 4140 steel*, Materiali in Tehnologije, 2015. **49**(5): p. 665-672.
 28. Günen, A., Karahan, İ.H., Karakaş, M.S., Kurt, B., Kanca, Y., Çay, V.V. and Yıldız, M., *Properties and corrosion resistance of AISI H13 hot-work tool steel with borided B₄C*, Metals and Materials International, 2019. **26**(11): p. 1-12.
 29. Krelling, A.P., Milan, J.C.G. and Costa, C.E., *Tribological behaviour of borided H13 steel with different boriding agents*, Surface Engineering, 2015. **31**(8): p. 581-587.
 30. Selcuk, B., Ipek, R. and Karamis, M.B., *A study on friction and wear behaviour of carburized, carbonitrided and borided AISI 1020 and 5115 steels*, Journal of Materials Processing Technology, 2003. **141**(2): p. 189–196.



Research Article

Experimental investigation and numerical verification of Coanda effect on curved surfaces using co-flow thrust vectoring

Emre Kara ^{a,*}  and Hüdai Erpulat ^a 

^aAeronautics and Aerospace Engineering Department, Gaziantep University, Faculty of Aeronautics and Aerospace, Gaziantep 27310, Turkey

ARTICLE INFO

Article history:

Received 29 June 2020

Revised 09 September 2020

Accepted 01 October 2020

Keywords:

Coanda effect

Computational fluid dynamics

Co-flow thrust vectoring

Numerical verification

ABSTRACT

In this study, a popular co-flow thrust vectoring system, which is superior to typical Coanda nozzles with one main jet, is examined experimentally and compared with 2D and 3D computational fluid dynamics results. High Speed Orienting Momentum with Enhanced Reversibility nozzle concept is the base design to proposed configuration which uses a control jet additional to the main jet for better and active enhancement on the flow vectoring and streamlined side-walls resulted in less flow blockage. This comparatively novel concept is utilized in an experimental setup to direct the thrust of aerial vehicles. The system includes two inlets (inlet1, inlet2) with different jet velocities and one pintle to separate and smoothly direct these jets and a converging-diverging nozzle to enclose these components. Experimental study is accomplished with four different configurations of inlet1 and inlet2 as 15 m/s and 10 m/s; 20 m/s and 10 m/s; 30 m/s and 10 m/s, and 45 m/s and 10 m/s, respectively. The tangential velocities on the curved surfaces are successfully measured utilizing a micro-manometer (Pitot tube) so that attachments/detachments of jets on the exit walls and deflection angles are calculated for each inlet velocities. The current experimental study also revealed that 3D assumption of computational fluid dynamics of Coanda effect is highly accurate and deflection angle results are not far from experimental results with the average deficit of only 5.44 %. As the result, 3D verification study resembles to experimental study in terms of deflection angles for all configurations.

© 2021, Advanced Researches and Engineering Journal (IAREJ) and the Author(s).

1. Introduction

This paper involves an experimental investigation of the Coanda effect thrust vectoring method in addition to a discussion about its potential implementation on aviation industry considering both its advantages and disadvantages. High Speed Orienting Momentum with Enhanced Reversibility (HOMER) nozzle concept [1] is selected as the base design to the proposed configuration. HOMER design is selected since it uses a control jet additional to the main jet for better and active enhancement on the flow vectoring and streamlined side-walls resulted in less flow blockage. The results of the deflection effect on the Coanda surfaces are also given in this paper.

Although these propulsion systems are developed for combat aircrafts, there is also a significant opportunity to

use them in the civil aviation sector. All those propulsion systems have advantages and disadvantages with respect to each other. In today's aviation world, the most commonly used thrust vectoring types are given in Figure 1. Among them Coanda methods [2, 3] are most suitable for new generation Vertical Take-Off and Landing/ Short Takeoff and Landing (VTOL/STOL) air vehicles, with their quick responses they can have a huge impact even on futuristic city transportation systems in the following decades.

In the field of aerial jet propulsion, it is crucial to correctly orient the thrust for take-off/landing accuracy and maneuverability. Almost all aeronautical vehicles rely on movement of aerodynamic control surfaces. Despite its reliability, control surfaces have plenty of heavy mechanisms.

* Corresponding author. Tel.: +90-342-317-3517 .

E-mail addresses: emrekara@gantep.edu.tr (E. Kara), erpulathudai06@gmail.com (H. Erpulat)

ORCID: 0000-0002-9282-5805 (E. Kara), 0000-0002-5709-7689 (H. Erpulat)

DOI: 10.35860/iarej.758397

This article is licensed under the CC BY-NC 4.0 International License (<https://creativecommons.org/licenses/by-nc/4.0/>).

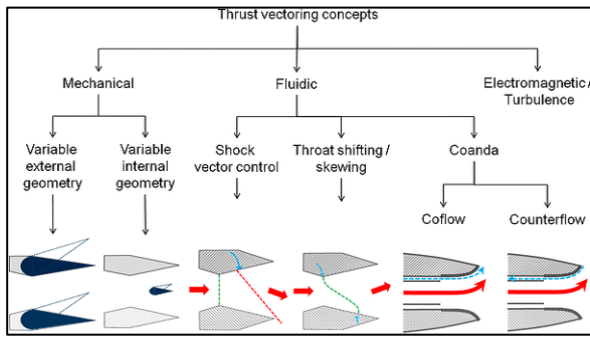


Figure 1. Thrust vectoring concepts [3]

In the last decade, the fluid thrust vectoring (FTV) concept is emerging as a popular choice [1-4], especially for the requirements of V/STOL aircrafts. The FTV method ensures the vectoring of the thrust without any movable mechanical parts simply and effectively. The affecting parameters of FTV are shortly classified as structure and flow parameters. Structure parameters include the nozzle geometry, location of the pintle tip, throat width. Flow parameters include the incoming jet(s), exit pressure, viscous flow over the smooth surface of the converging-diverging nozzle. In traditional nozzle systems, the flow has only one directed incoming jet and is oriented by the movement of exit diverging nozzle walls [4].

Coanda effect is the adhesion of a high speed jet to a convex surface. Coanda effect is essentially utilized in aerospace applications [5], heating/cooling [6] and marine technology [7]. There are three alternative nozzle configurations for the utilization of Coanda surfaces: Newman setup [4] with only one jet next to the convex surface, Juvet setup [8] with two primary jets and a control jet and HOMER nozzle setup [1] composed of two inlets, two Coanda surfaces, one pintle to separate and smoothly direct these jets and a converging-diverging nozzle to enclose whole components. In the patented study of 2D nozzle geometry, this nozzle concept is developed by Trancossi et al. [9] who finished an EU project (ACHEON) [1] regarding this exciting design. In the current study, it is aimed to develop a FTV concept by using similar Coanda surfaces. HOMER nozzle concept [1] is selected as the base design to the proposed configuration since it uses a control jet additional to the main jet for better and active enhancement on the flow vectoring and streamlined side-walls resulted in less flow blockage. Co-flow FTV is mostly used one in subsonic flows having one constant velocity and other jet stream with a controllable velocity is given as red and blue streams, respectively, as shown in Figure 2. This differential velocity allows the thrust vector slope becoming more controllable.

Although general purpose fluidic thrust vectoring is utilized for enhancing thrust maneuverability since late 1960s [10], the HOMER design is highly immature and

open to improvement with its only 10 years of background. In this paragraph, the most recent article/project of the study groups working on HOMER design is taken into account. Starting from the original design of Trancossi et al. [1], Subash and Dumas [11] represented different turbulence models applied on air-jet flow tangential to a curved surface. They realized that jet deflection angle and the thrust are important parameters that should be handled with extreme care. They do not present an experimental study to compare and validate their CFD results. Another study group, Cen et al. [12] tried to integrate HOMER design to Unmanned Aerial Vehicle (UAV) for enhancing its maneuverability. They focused on developing this design starting from CFD simulation of the design, they proposed an integrated flight/thrust vectoring (HOMER) control scheme for fixed wing UAVs. They stated the effectiveness of the HOMER design adding to four-cascaded nonlinear dynamic inversion (NDI) control law they proposed. However, they do not build an experimental setup as in the previous reference to validate their results. In 2016, Trancossi et al. [13] generated the mathematical model of HOMER nozzle. Although the mathematical model of 2D case of the system is developed in the article, it can be counted as a step stone for future 3D mathematical modeling attempts. They validate the model by comparing the results with CFD simulations of velocity profiles and pressure drops through the selected sections. They found a good correlation of those outcomes despite frictional effect underestimation. In the most recent study, Panneer and Thiyagu [14] analyzed effect of pintle geometry on HOMER design. They found that sharper tip of a pintle can result in an increase on deflection angle. Additionally, the velocity plots are shown for different velocity ratio configurations. Although this newest study shows a new perspective on HOMER nozzle effectiveness, CFD results are not enough to confirm the HOMER nozzle concept and a new viewpoint should be developed both in experiments and numerical approaches. In the current study, both 3D CFD study and construction of a new HOMER nozzle design, experimental study and the verification of the results using the CFD study are all served to literature as one original complementary package.

HOMER nozzle concept [1] basically directs the jet direction as velocity of the main and control jet streams changes and clearance between the upper and lower walls of Coanda surface optimized. This configuration enhances the aircraft thrust system in following steps:

- Thrust vectoring capability increases with avoided stall on the diverging section of each wall.
- No-mechanical-part in movement results in reduction in total weight and total cost of structure.
- Quick and directional control of propulsion may be used as an advantageous maneuverability of V/STOL

aircrafts.

- Using more efficient control system on thrust vectoring means less use of fossil-based fuel.

2. CFD Study

In the previously published study [15], commercially available CFD software, ANSYS 17.0, was used as the mesh generator (ANSYS Meshing) and the flow solver (ANSYS Fluent). The boundary conditions for the solution domain are shortly depicted in Figure 3 both for previous 2D study and the currently done 3D study. Same parameters are employed for 3D study to compare them independently so that one can refer to previous study.

As suggested in [11], velocity ratio ($VR = V_2/V_1$) between channels inlet1 and inlet2 were selected as ranging from 1.5 to 3.0, i.e. 15 m/s and 10 m/s; 20 m/s and 10 m/s; 30 m/s and 10 m/s, respectively. In order to understand whether the upper limit of the velocity ratio is correctly stated as 3.0 in reference [11], a new case study with inlet channel velocities of 45 m/s and 10 m/s is also studied. Details can be found in Kara and Erpulat [15]. In the current experimental study, this previously worked CFD results are also referenced and compared at the end of the next section in addition to 3D CFD study that is shown to be more effective in the prediction of deflection angle over the Coanda surface.

As seen in Figure 4, four different case studies are selected for working fluid (ideal gas as air) with inlet1 (constant) velocity of 10 m/s and four different inlet2 (control) velocities of 15, 20, 30 and 45 m/s for case study 1, case study 2, case study 3 and case study 4, respectively. Deflection angle is increased from 50° to a physical limit of 180° when VR increased from 1.5 to 4.5 respectively. These results are compared with experimental study and are also discussed in the proceeding section.

3. Experimental Study

The experimental setup consists of two Coanda surfaces, one pintle, two convergent ducts to direct the flow through the fans as shown in Figure 5. All the designed parts are produced utilizing 3D printer with PLA+ material.

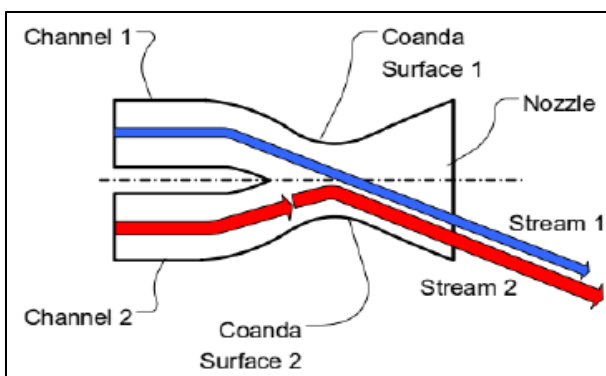


Figure 2. Schematic of the HOMER nozzle [2]

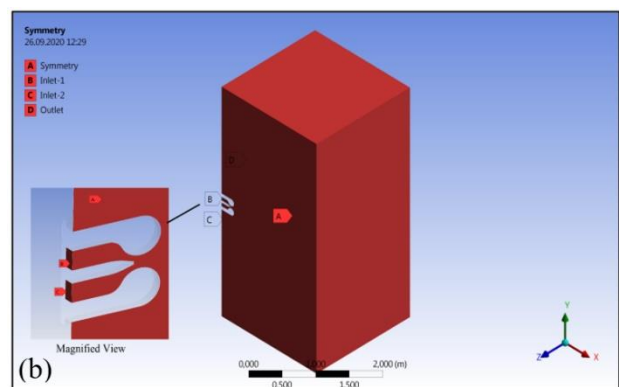
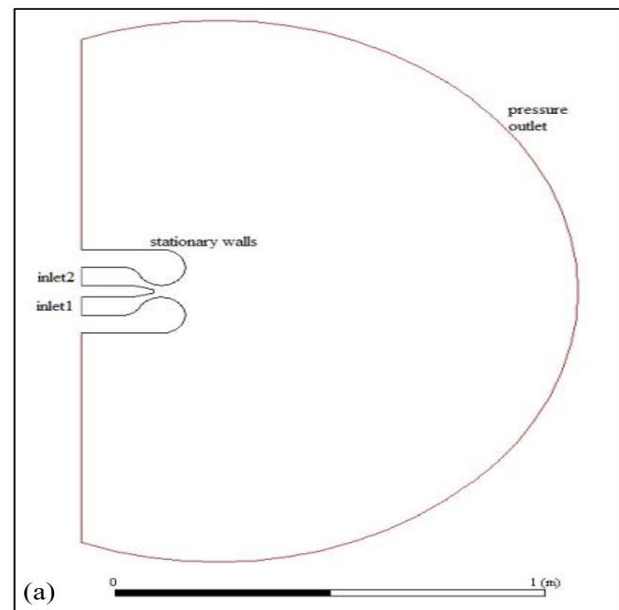


Figure 3. Boundary conditions (velocity inlets, pressure outlet, symmetry plane) of suggested FTV design (a) 2D, (b) 3D

Application page of the 3D printer is captured and is shown in Figure 6. Propellers with ten blades are preferred due to the rotational speed of the selected motor is not enough for less blade propeller as empirically tested several times. DC motor is selected with 1400 kV, 27.8 mm x 27.8 mm duct area and 3.17 mm shaft diameter.

Plexiglass is used as cover material on the side surfaces as shown in Figure 5(a). This material is quite cheap and easy to manage. Coanda surfaces are completely covered with plexiglass. The ducted fans are combined appropriately by plexiglass support. The inner corner of the Coanda surfaces is attached to the side surfaces by means of screws. This connection allows Coanda surfaces to move but remain in the desired position. The transducer surfaces are fastened to the Coanda surfaces by side screws. Motor cables are pulled out from the holes which drilled on the circle to rectangle converter and connected to the motor drives. PCE-PFM 2 brand Micro Manometer with Pitot tube (PCE Instruments, Southampton, UK) is employed for accurately measuring air flow speed in m/s with PCE software for data processing. List of equipment mentioned here are summarized in Table 1 with their working specifications.

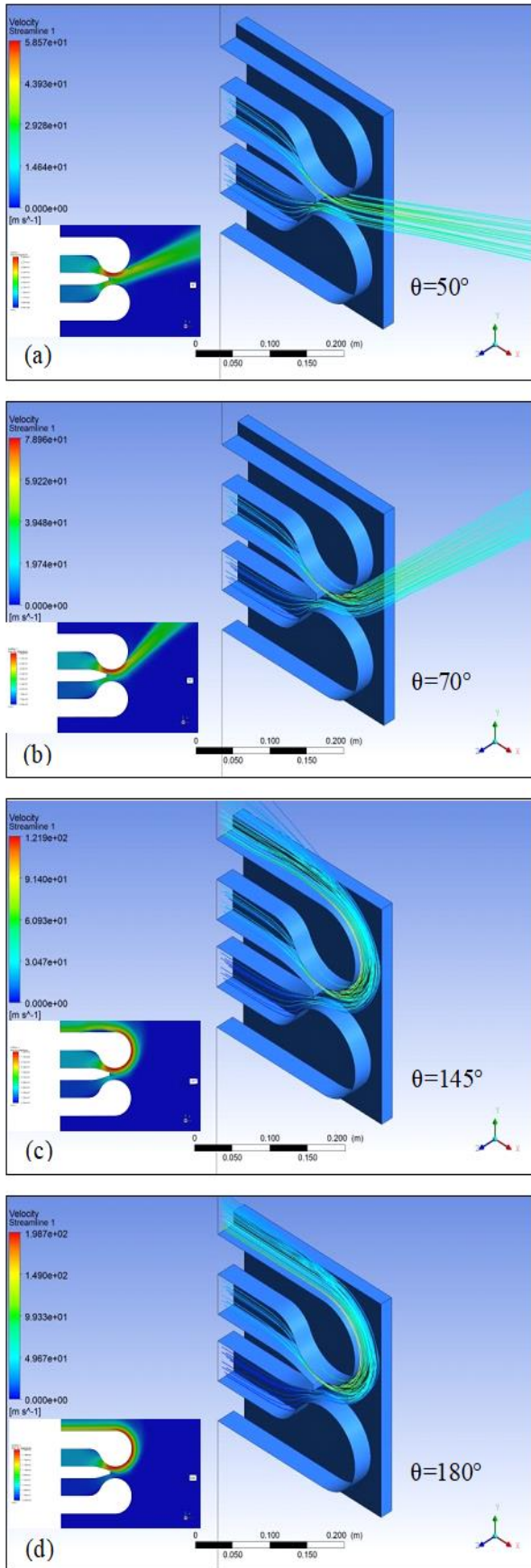


Figure 4. 3D streamlines and velocity contours projected onto symmetry surface for VR of (a) 1.5, (b) 2.0, (c) 3.0 and (d) 4.5 with deflection angles (θ) next to them.

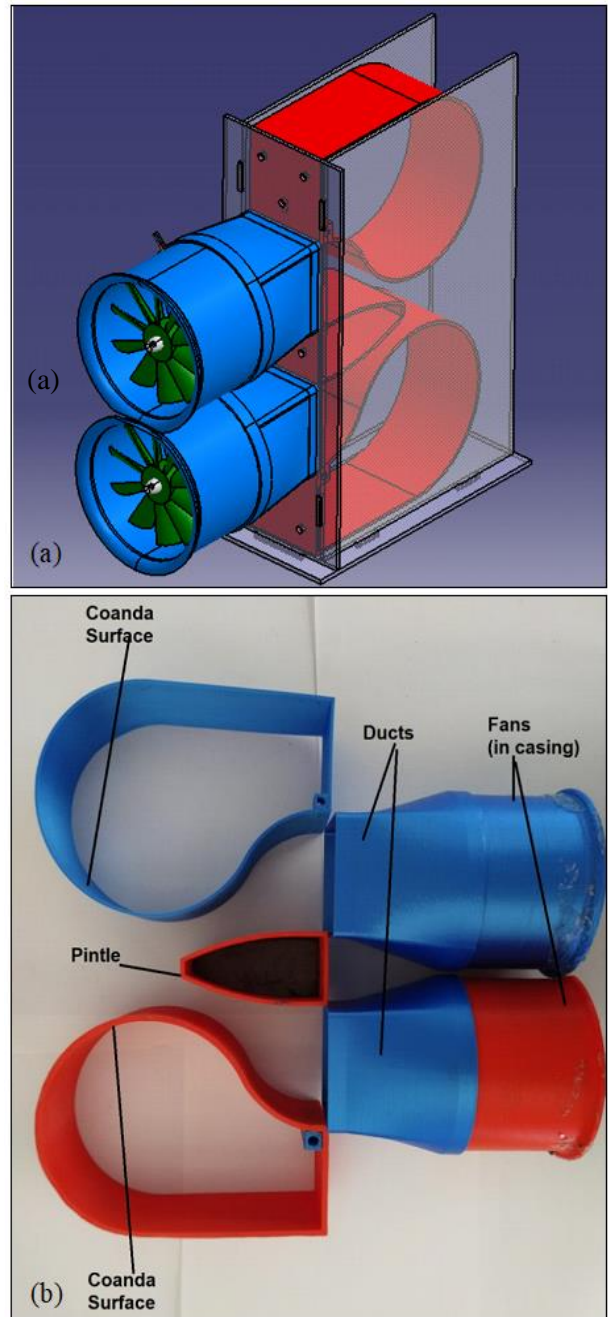


Figure 5. (a) 3D model of the experimental setup and (b) the constructed experimental setup with components explained

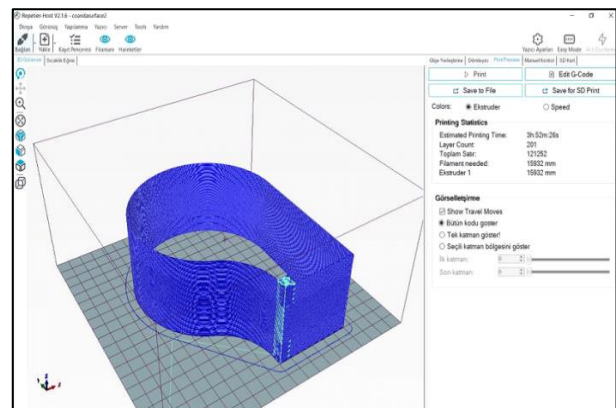


Figure 6. Repetier-host software snapshot of utilized 3D printer

Table 1. Used equipment and their working specifications

Brand name	Equipment name (quantity)	Specifications
PCE-PFM 2	Micro manometer	For pressure: M. ranges: 50 mbar Resolution: 0.01 mbar Uncertainty: $\pm 0.3\%$ of the measuring range Max. pressure: 10 psi Response time: 0.5 seconds For air velocity: Measurement ranges: 1...80 m/s Resolution: 0.01 m/s Uncertainty: $\pm 2.5\%$ of the measuring range Dimensions: Unit: 210 x 75 x 50 mm Pitot tube: 375 x 205 mm Weight: 280 g
Wester A2212	DC motor (2)	Brushless outrunner motor: Shaft diameter: 3.17 mm Dimensions: 27.8 mm x 27.8 mm Recommended Propeller: 2-cells 7 x 3.5 or 3-cells 6x3 Weight: 48g

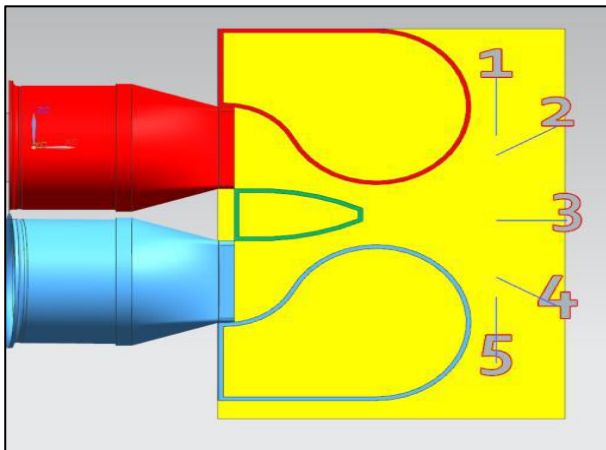


Figure 7. Velocity measurement sections for micro manometer with Pitot tube

Table 2. Velocity Ratio versus sections of velocity measurements over the Coanda surfaces

VR	Section 1	Section 2	Section 3	Section 4	Section 5
1.0*	2.1 m/s	4.3 m/s	12.2m/s	4.3 m/s	2.1 m/s
1.5	6.7 m/s	7.1 m/s	6.0 m/s	-1.2m/s	-0.4 m/s
2.0	9.9 m/s	8.2 m/s	6.9 m/s	-1.8m/s	-0.8 m/s
3.0	9.2 m/s	4.8 m/s	3.0 m/s	-2.0m/s	-1.0 m/s
4.5	13.1m/s	6.1 m/s	3.4 m/s	-2.8m/s	-1.2 m/s

*Baseline control test case

The measured data are transferred to a computer with the included USB adapter and software. Measurement locations are shown in Figure 7. Sections are separated 45° apart from each other and baseline is taken as section 3 (the horizontal section). All results are tabulated in Table 2.

As expected, the greater the difference between the electrical motors speeds means the more vectoring of the co-flow. Furthermore, during the execution of the experiments, it was observed that the Coanda effect is increased when the surfaces were brought closer to each other and the adhesion effect increased when the flow was forced through a narrower opening over the pintle. Considering experimental results of VR's for different sections in Table 2, curves are fitted for each set of experiments with R^2 of 0.96 or more. Thus, Table 3 summarizes the experimental outcomes of deflection angles at different VR's.

Comparing previously studied 2D CFD study [15] with the current experimental study reveals that 2D assumption of the fluid dynamics of Coanda effect is highly inaccurate. Moreover, it overshoots the deflection angle, θ , in low-to-moderate VR values and after $VR > 2.0$ it changes direction and underestimates the deflection angle, θ , at a total absolute average of 20.34 %. As seen in the 4th and 6th columns of Table 3, 3D CFD study is more accurate and is not far from experimental results with the absolute average deficit of only 5.44 %. This can be the result of secondary flows in third dimension and effect of side surface guidance which is not the case for a 2D study.

Authors understood that an assumption of no secondary flow and dominantly axisymmetric behavior of the flow are both incorrect. Thus, it is studied in three-dimensional computational domain with careful statement of the flow behavior in boundary layer and separation regions over the Coanda surfaces. Experimental outcomes of the current work are carefully examined and compared with 3D CFD study. In order to understand whether the upper limit of the velocity ratio is correctly stated as 3.0 in reference [11], a new case study with inlet channel velocities of 45 m/s and 10 m/s is also studied experimentally as shown in Table 2 and the limiting VR is found to be 4.5 in the experimental setup as in the cases of CFD studies.

Table 3. Velocity Ratio versus experimental deflection angles, θ , (with $R^2 \geq 0.96$) and numerical (CFD) deflection angles, θ

VR	θ (Exp)	θ (2D CFD)	θ (3D CFD)	Error of 2D CFD results [%]	Error of 3D CFD results [%]
1.0*	0.0	0	0	0.00	0.00
1.5	56.8	65	50	14.44	-11.97
2.0	66.4	80	70	20.48	5.42
3.0	134.0	95	145	-29.10	8.20
4.5	177.0	110	180	-37.85	1.69
Average error relative to experiments				20.37	5.44

*Baseline control test case

4. Conclusion

In this study, a novel propulsion system is examined. It is understood how this kind of effort can change future of the thrust vectoring in VTOLs.

As one of the key outcomes, firstly, it is observed that the velocities of the air supplied from two different ducts are controlled free from each other and this concept is proved experimentally. Throughout the numerical study, the effect of geometry properties on mesh properties was examined in 3D. Particularly, the importance of pre-study of experiments was realized utilizing the desired quality of the mesh. Results revealed that the outcomes of the setup are not compatible with 2D CFD results, especially in the range. This is because two-dimensional approach for fluid dynamics of the Coanda effect is not adequate to correctly define the real flow as explained in the results section. Thus, the experimental results are compared with 3D CFD analysis considering secondary flows that are not possible with 2D approach. Results showed that 3D verification study resembles to experimental study in terms of deflection angles for all configurations.

The new insights that should be investigated and be revealed can be summarized as follows:

- Experimental setup should be improved after currently in-progress project so that powerful motor in ducted fans, superior materials for the cover without unwanted airflow leakages, etc.
- Visualization will be added as another improvement to the current experimental study employing Schlieren flow observation techniques.
- Better and more expensive choices will be investigated for flush-mounted shear stress sensors and static pressure measurement devices in order to increase the diversity of focused techniques. Thus, the behavior of the airflow over Coanda surfaces will be understood and inquired profoundly.

Declaration

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article. The authors also declared that this article is original, was prepared in accordance with international publication and research ethics, and ethical committee permission or any special permission is not required.

Author Contributions

E. Kara and H. Erpulat conceived the study together. E. Kara supervised the research, analyzed the data and drafted / finalized the paper. H. Erpulat conducted the experiments, contributed to the data analysis and revised the paper.

Nomenclature

θ	: Deflection Angle [°]
FTV	: Fluid Thrust Vectoring
CFD	: Computational Fluid Dynamics
HOMER	: High Speed Orienting Momentum with Enhanced Reversibility
NDI	: Nonlinear Dynamic Inversion
STOL	: Short Takeoff and Landing
UAV	: Unmanned Aerial Vehicle
VR	: Velocity Ratio
VTOL	: Vertical Take-Off and Landing

References

1. Trancossi, M. and A. Dumas, ACHEON: *Aerial Coanda High Efficiency Orienting-jet Nozzle*. SAE Technical Paper, 2011. No. 2011-01-2737.
2. Trancossi, M., A. Dumas, S. S. Das, and J. Pascoa, *Design methods of Coanda effect nozzle with two streams*. Incas Bulletin, 2014. 6(1): p. 83-95.
3. Bougas, L., and M. Hornung, *Propulsion system integration and thrust vectoring aspects for scaled jet UAVs*. CEAS Aeronautical Journal, 2013. 4(3): p. 327-343.
4. Newman, B. G., *The Deflexion of Plane Jet by Adjacent Boundaries Coanda Effect*. 1961, UK: Pergamon Press.
5. Jain, S., S. Roy, D. Gupta, V. Kumar, and N. Kumar, *Study on fluidic thrust vectoring techniques for application in V/STOL aircrafts*. SAE Technical Paper, 2015. No. 2015-01-2423.
6. Sidiropoulos, V., and J Vlachopoulos, *An investigation of Venturi and Coanda effects in blown film cooling*. International Polymer Processing, 2000. 15(1): p. 40-45.
7. El Halal, Y., C. H. Marques, L. A. Rocha, L. A. Isoldi, R. D. L. Lemos, C. Fragassa, and E. D. dos Santos, *Numerical study of turbulent air and water flows in a nozzle based on the Coanda effect*. Journal of Marine Science and Engineering, 2019. 7(2): 21.
8. Juvet, P. J. D., *Control of high Reynolds number round jets*, in Mechanical Engineering 1993, Stanford University: USA, TF-59.
9. Trancossi, M., A. Dumas, I. Giuliani, and I. Baffigi, *Ugello Capace di Deviare in Modo Dinamico e Controllabile un getto Sintetico senza parti Meccaniche in Movimento e suo Sistema di Con trollo*, 2011, Patent No. RE2011A000049, Italy.
10. Springer, A., *50 Years of NASA Aeronautics Achievements*. 46th AIAA Aerospace Sciences Meeting and Exhibit , p. 859.
11. Subhash, M., and A. Dumas, *Computational study of Coanda adhesion over curved surface*. SAE International Journal of Aerospace, 2013. 6(2013-01-2302): p. 260-272.
12. Cen, Z., T. Smith, P. Stewart, and J. Stewart, *Integrated flight/thrust vectoring control for jet-powered unmanned aerial vehicles with ACHEON propulsion*. Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering, 2015. 229(6): p. 1057-1075.
13. Trancossi, M., J. Stewart, S. Maharshi and D. Angeli, *Mathematical model of a constructal Coanda effect nozzle*. Journal of Applied Fluid Mechanics, 2016. 9(6): p. 2813-2822.
14. Panneer, M., and R. Thiyagu, *Design and analysis of Coanda effect nozzle with two independent streams*. International Journal of Ambient Energy, 2020. 41(8): p.


851-860.

15. Kara E., and H. E., *Numerical Investigation of Jet Orientation Using Co-Flow Thrust Vectoring with Coanda Effect*, in ICAME2019: İstanbul, p. 1-8.



e-ISSN: 2618-575X

INTERNATIONAL ADVANCED RESEARCHES
and
ENGINEERING JOURNAL

Journal homepage: www.dergipark.org.tr/en/pub/iarejInternational
Open Access Volume 05
Issue 01

April, 2021

Research Article

Modelling of cutting parameters for Nilo 36 superalloy with machine learning methods and developing an interactive interface

Gültekin Basmacı ^{a,*} , İsmail Kirbaş ^b  and Mustafa Ay ^c 

^aDepartment of Mechanical Engineering; Burdur Mehmet Akif Arsoy University, 15030, Burdur, Turkey

^bDepartment of Computer Engineering; Burdur Mehmet Akif Arsoy University, 15030, Burdur, Turkey

^cDepartment of Mechanical Engineering; Marmara University, 34000, Istanbul,

ARTICLE INFO*Article history:*

Received 04 October 2020

Revised 22 January 2021

Accepted 28 January 2021

Keywords:

Cutting forces

Interface development

Machine learning

Nilo 36 superalloy

Response surface design

ABSTRACT

Superalloys have become increasingly used in the machining sector due to their high strength, temperature and machinability. One of these alloys, Nilo (Invar) 36, has a low thermal expansion and its use is rapidly increasing in areas where high temperature and expansion are not required, especially in composite mould applications, such as aerospace, electronics, measuring instruments and aerospace. In this study, a mathematical model based on artificial intelligence and an interactive visual interface in MATLAB software were developed according to the test results obtained from surface roughness Ra, cutting methods, rotational speeds, cooling method and cutting speed of Nilo 36 alloy. For the mathematical analysis of the measurements, the number of experiments to be performed by using Minitab program and Taguchi method was reduced to 32. The measurement results were modelled by Response Surface Design method and the factors affecting the surface roughness were determined in order of importance. A high-performance feed-forward artificial neural network has been developed using experimental data and an interactive interface has been prepared based on the developed model. Thus, the user can easily observe the cutting forces and surface roughness values for different cutting parameters with high accuracy.

© 2021, Advanced Researches and Engineering Journal (IAREJ) and the Author(s).

1. Introduction

Invar 36 (Nilo 36) is a 36% Nickel and 64% iron alloy and is the material with the lowest thermal expansion between the metals and alloys in the temperature range of 25-230°C. Ha and Min claim that Invar 36 is an alloy that combines high strength, stiffness, thermal and dimensional stability Tae and Seok [1]. It is used in laser equipment, control devices, reference measuring devices, gas cookers and heaters with low expansion feature over a wide temperature range. It is also an important material for aerospace, telecommunications, cryogenic tanks, composite mould production Tae and Seok [1].

Maranhão and Davim showed that Invar 36 material is relatively difficult to process due to its properties. High cutting temperature, high cutting forces, rapid tool wear, chip breakage, chip adherence cause decreased surface quality Cesar et al. [2]. In the turning process, heat energy

is generated in the plastic deformation process of the work-piece, mostly in the first deformation zone. Tekaslan et al. [3] have shown that the process efficiency and tool life have been directly affected by cutting length and tool geometry Tekaslan et al. [3].

It is seen that many experimental studies have been performed for investigating the effects of various parameters used on the processing of similar materials on roughness of surface and cutting forces. Li et al. [4]. evaluated the deviations between the theoretical and experimental results for the cutting forces in the turning process of AISI 304 austenitic stainless steel with TiC coated cutting tool and emphasized that the theoretical approach can be used with 80% accuracy. Diniz et al. [5] indicated the machinability of AISI 304 steels with coated cemented carbide cutting tools has been performed and the cutting speed is an important parameter for surface

* Corresponding author. Tel.: +90 248-213-2700; Fax: +90-248-213-2704

E-mail addresses: gbasmaci@mehmetakif.edu.tr (G.Basmacı), ismailkirbas@mehmetakif.edu.tr (I.Kirbas), muay@marmara.edu.tr (M.Ay)

ORCID: 0000-0003-4818-3160 (G.Basmacı), 0000-0002-1206-8294 (I.Kirbas), 0000-0002-7672-1846 (M. Ay)

DOI: 10.35860/iarej.805124

This article is licensed under the CC BY-NC 4.0 International License (<https://creativecommons.org/licenses/by-nc/4.0/>).

roughness (Ra). Kaladhar et al. [6]. Reported that when turning AISI 304 stainless steel with low feed rate and high cutting speed, it was stated that the sound pressure level during cutting decreases and thus Ra decreases. Dirviyam and Palanisamy [7] claim that the use of a cutting tool with wiper geometry during turning has a positive effect on surface quality. Basmacı et al. [8]. found out that the use of a wiper tip eliminates the second process, reducing process steps and time. However, the increase of the wiper surface causes the cutting forces to increase Ay [8]. In the study of parameter optimization in the process of turning AISI 304L stainless steel, the wiper tip had a positive effect on surface quality.

Basmacı and Ay [9]. Basmacı et al. [10], emphasized that production quantity, economy and time are not sufficient for a manufacturing process Dhar et al. [11]. In addition, it is directly related to the harm to the environment and its impact on human health. There are studies using different cooling techniques sensitive to human health and environmental pollution in turning operations. It has been observed that the performance of these alternative cooling techniques is superior to the general cooling technique. Dhar et al. [11] reported that MQL cutting speed, feed rate and cutting temperature have important contributions to surface integrity. Itoigawa et al. [12]. reported that MQL reduces friction between the tool and the workpiece. Sheng et al. [13]. emphasized that the main effect of MQL is to lower the cutting temperature. Marco et al. [14] reported that MQL decreased the formation of crater wear Basmacı et al. [15]. showed that the MQL technique has a positive result on the reduction of heat and wear between tool-chip and workpiece-chip and in the cutting zone, as well as on surface quality improvement Basmacı et al. [15].

In this study, the effect of cutting speed, cooling system and cutting method on surface roughness and cutting forces in turning of Nilo 36 steels were investigated. Taguchi technique has been used to determine and optimize the effect of cutting speed, cooling system and cutting method of Nilo 36 steels on turning surface roughness and cutting forces.

In addition, analysis of variance (ANOVA) was performed to figure out the effect of each parameter on the results and the optimum machinability of Nilo 36 steel. A full autonomous design of proving was formulated in this paper to study a long- term process coefficient for both materials. The bulk density, tensile mechanical attributes, fractography, material composition, and remaining stresses of the parts produced were investigated.

An optimum process window has been proposed based on experimental work Yakout et al. [16]. The most significant contribution of the study to literature is the artificial intelligence supported interactive interface developed within the scope of the study. Using this

interface, it is possible to predict with high accuracy what size F forces will be encountered during the process and what the surface roughness value will be according to various cutting parameter values and different cutting methods.

Normally, in order to know these values, very valuable and expensive materials must be consumed and repeated trials are required. In addition, there is a loss of labour, energy and time. Thanks to the implemented interface, the user can manually change the cutting parameters, cutting methods, cutting speed and type and the amount of coolant, and the result of the process of tens of combinations can be calculated instantly and displayed graphically.

2. Materials and Methods

2.1 Materials

In the present study, Nilo36 stainless steel with 150 mm length and 50 mm diameter was used as test material and its chemical characteristics are given in Table 1.

2.2 Experimental Design and RSM Analysis

In the presented study, Nilo36 alloy was processed in four different cutting speed (100, 140, 180, 220 rpm), two different cutting methods (conventional and wiper) and four different cooling methods (dry cut, MQL 20ml, MQL 40ml and CO₂). According to the experimental results, a model based on artificial neural network and also a visual interface was developed in MATLAB software environment.

In the experiments, CNMG 12 04 08 MM SUMITOMO cutting tips, Johnford TC 35 CNC Fanuc OT an x-z axis CNC and Mahr perhometer/M1 type surface roughness meter was used. In addition, Kistler 9121 dynamometer system together with a Kistler 5019 charge amplifier and DynoWare software were utilized for the force determination (Figure 1.)

Table 1. Nominal Compositions, %

Alloy	Ni	Fe	Others
NILO alloy 36	36.0	64.0	-
Density			
Alloy	g/cm ³	lb/in ³	
NILO alloy 36	8.11	0.293	
Thermal Conductivity at 20°C (68°F)			
Alloy	W/m°C	Btu in/ft ² h °F	
NILO alloy 36	10.0	69.3	
Electrical Resistivity			
Temperature (Microhm cm (ohm.circ mil/ft)			
°C	°F	NILO alloy 36	
20	68	80 (481)	
Typical Mechanical Properties of NILO alloy 36			
Yield Strength (0.2% Offset)		Tensile Strength	
MPa	ksi	MPa	ksi
240	35.0	490	71.0
Elastic Modulus Data			
Elastic Modulus			
Alloy	GPa	10 ³ ksi	
NILO alloy 36	140	20.3	



Figure 1. Experimental setup

Figure 2 demonstrates the process flow performed within the scope of the study in detail. Invar alloy with an elastic modulus effect over a wide temperature range. Intensive microscopic characterizations indicate the occurrence of a prematurity transformation with formation of nano-scale domains that can be restricted by deformation Qin et al. [13].

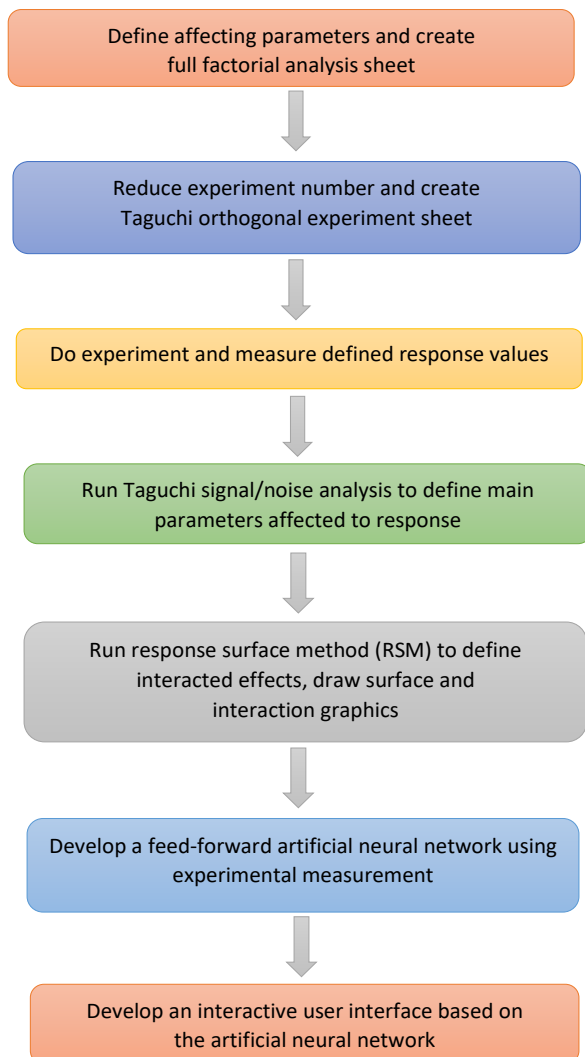


Figure 2. The process flow performed during the study

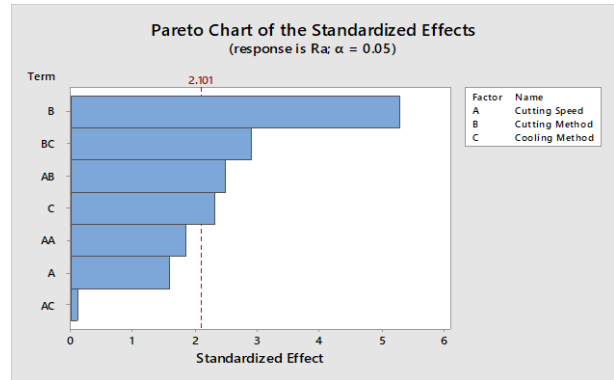


Figure 3. Pareto chart of the standardized effects on Ra

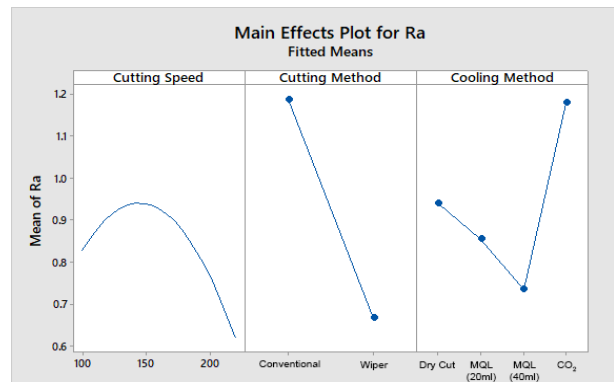


Figure 4. Main effects plot for Ra

First of all, Minitab program was used for mathematical analysis of the measurements and the number of experiments was reduced to 32 using Taguchi method. The measurement results were modelled by Response Surface Design method and the factors affecting the surface roughness were determined in order of importance. ANOVA calculation based on cutting speed, cutting method and cooling method factors is given below.

The coefficient of multiple determination for multiple regression (R squared) value was calculated as 78.88%. This value shows that the mathematical model is acceptable. Figure 3 shows the influencing parameters on the surface roughness and their degree of influence as a pareto graph. When standardized factors are compared, the most important factor affecting the surface roughness has been the cutting method. Following the method of cutting method cooling method pair, cutting speed cutting method and finally cooling method stand out as effective parameters. Figure 4 shows how the cutting speed, cutting method and cooling method parameters affect the mean Ra value.

When Figure 4 is examined, higher values for cutting speed cause lower surface roughness, wiper should be selected for cutting method and MQL 40 should be preferred for cooling method. The use of CO₂ as a cooling method gives negative results in terms of surface roughness compared to other cooling methods. There is also a difference in surface roughness between the conventional cutting method and the wiper, and the wiper

method proved to be more successful.

RSM is a mathematical modelling and statistical assessment method used to optimize a cutting system's input parameters. This method is used to create multi-factor models, obtaining quantitative data from a suitable experimental design. Such models can be displayed graphically. In graphs, a response surface is used to evaluate how factors affect response, to explain the relationship between variables, and to show the combined effects of factors. Kirbaş et al. [17], Asilturk et al. [18].

RSM method consists of three stages. Firstly, physical experiments are carried out to obtain response values using experimental parameters combination. It means a decrease in costs in comparison to the number of tests performed using conventional methods with less and more efficient experiments. Searched intermediate response values could be determined in a short time. The relation between the input parameters and the reactions obtained is defined in the second stage as a quadratic or exponential polynomial. In the third stage, an analysis like ANOVA or surface graphics can determine the optimum points Abou-El-Hossein et al. [19]. Taguchi orthogonal array method provides requiring information by conducting limited experiments. Thus it, is time-energy-material saving method. The depth of cut is the heaviest coefficient that impacts the surface roughness, followed by the cutting speed and feed rate in case of milling of INVAR-36 Khanna et al. [20].

The correlation between parameters and their corresponding responses is typically given in RSM problems using the following quadratic polynomial Equation (1) Kirbas et al. [17].

$$\eta = \beta_0 + \sum_{i=1}^k \beta_i X_i + \sum_{i=1}^k \beta_{ii} X_i^2 + \sum_i \sum_j \beta_{ij} X_i X_j + \varepsilon \quad (1)$$

In the equation (1), η represents the estimated responses (Fx, Fy and Fz); β_0 is a constant; β_i and β_{ii} show the first and the second-degree encoded input parameters and β_{ij} are the coefficients of parameter interactions.

Figure 5 shows a 3-D surface plot showing the relationship between cutting method and cutting speed affecting Ra. For the cutting method on the X axis, 1 represents the conventional method and 2 represents the wiper method.

In Figure 6 and Figure 7, the cutting speed effect and cooling method factors on surface roughness are given as 3D surface graph and contour plot. For cooling method, 0 indicates dry cut and 100 indicates CO₂ cooling respectively.

During the machining of the Nilo36 alloy on the lathe, the forces Fx, Fy and Fz were also examined and a mathematical model was developed. In the ANOVA analysis given below, the factors affecting cutting speed,

cutting method and cooling method were used and its effects on Fx force were calculated.

The performance of ANOVA model was 88.20%. Accordingly, the most important factor affecting Fx force emerges as cutting method. Then comes the cooling - cutting method duo and cooling method. According to the results of the analysis, the effect of cutting speed factor is much less than other factors.

The factors affecting the Fx force and their degree of influence are clearly seen in the pareto graph given in Figure 8. The factors affecting Fx force were determined as cutting speed, cutting - cooling method duo and cooling method.

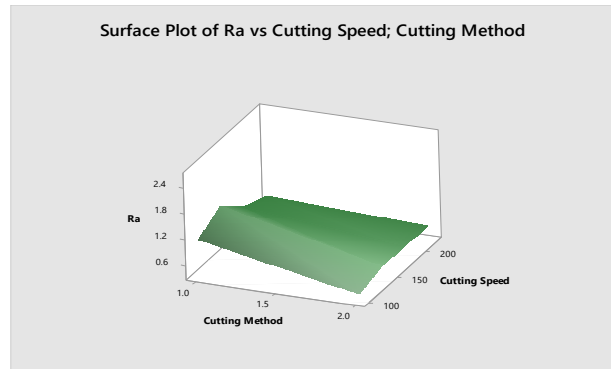


Figure 5. Surface plot of Ra vs cutting speed and cutting method

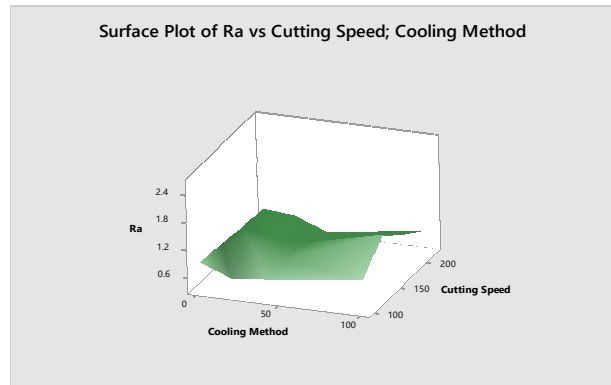


Figure 6. Surface plot of Ra vs cutting speed and cooling method

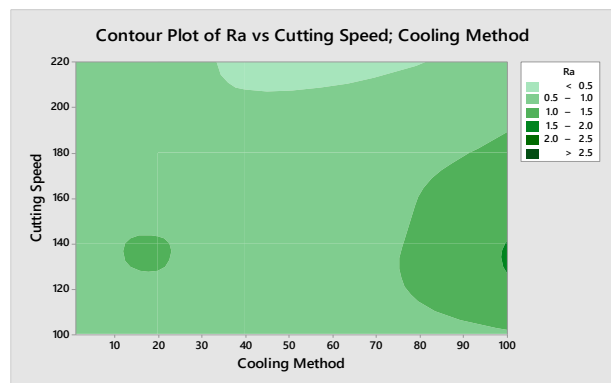


Figure 7. Contour plot of Ra vs cutting speed and cooling method

In Figure 9, factors and parameters affecting Fx force are given. It is seen that the lowest Fx values have been obtained at the lowest and highest cycles when conventional was selected as cutting method and MQL 40 ml was selected as cooling method. When using Wiper method, higher values of Fx force were measured compared to conventional method. When CO₂ was selected as cooling method, Fx force increases.

The results of ANOVA analysis to calculate the factors affecting Fy force are given below. The model performance value was found to be 77.01%, indicating that the model was acceptable.

Pareto analysis graph is given in Figure 8. The most decisive parameter for the Fy force is the cutting method. The cutting speed duo emerges as the second factor that crosses the pareto boundary.

Factors affecting Fx and Fy forces exhibit similar characteristics. The effect of the cooling method for Fy force is lower and uncertain than Fx. The effect of the cutting method is increased in the wiper method as in the Fx force and there is an obvious difference.

ANOVA analysis was performed to determine the factors affecting Fz force and 84.50% performance ratio was calculated for the developed model. This value is quite high and shows that the model is mathematically acceptable.

In the pareto graph given in Figure 11, it is obvious that the most obvious components for force Fz are the cooling method and the cutting method. However, the effect of cooling method was much higher than other effects.

It is seen that the lowest value for cooling method is obtained in MQL (40ml) process as in other force values. There is quite a difference between CO₂ and MQL method. As cutting speed value increases, Fz value decreases. Selecting conventional as cutting method leads to lower Fz force value than wiper method.

3. Artificial Neural Network Modelling and Developing an Interactive Interface

When the literature is examined, there are examples where cutting forces and cutting parameters are modelled with an artificial neural network [21-24]. However, none of these examples mention the development of the graphical user interface included in this study.

Ra, Fx, Fy and Fz values obtained by processing Nilo36 alloy were loaded into an artificial neural network consisting of 12 neurons and a total of 3 layers using MATLAB software.

The input parameters of the artificial neural network were determined as cutting speed, cutting method and cooling method. Ra, Fx, Fy and Fz values were estimated as output values. Figure 14 shows the input, output and latent layers of the trained neural network and the number of neurons in each layer.

Factors affecting the Fy force and parameter values are shown in Figure 10. The factors affecting Fz force and parameter values are given in Figure 12 and Figure 13.

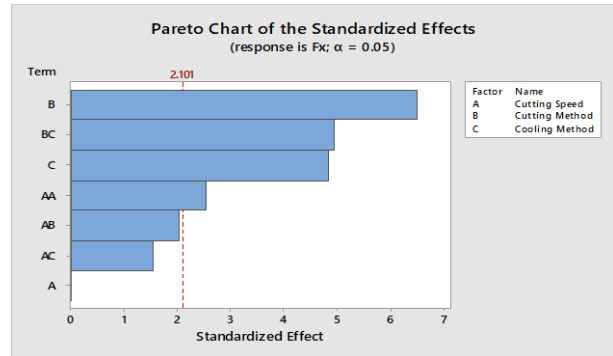


Figure 8. Pareto chart of the standardized effects on Fx force

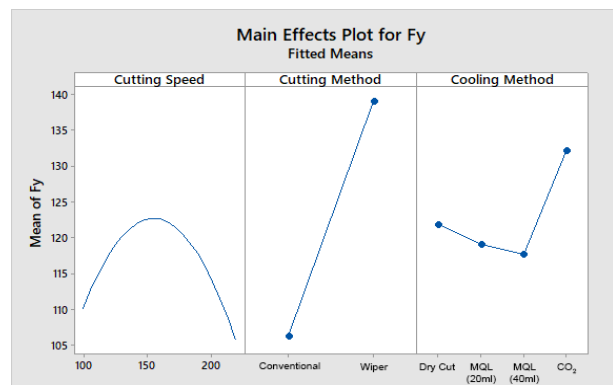


Figure 9. Main effects plot for Fx force

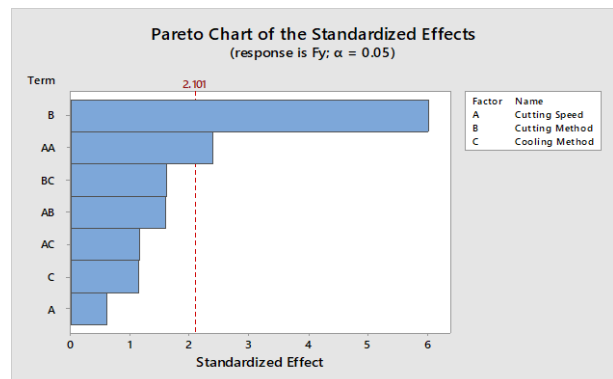


Figure 10. Pareto chart of the standardized effects on Fy force

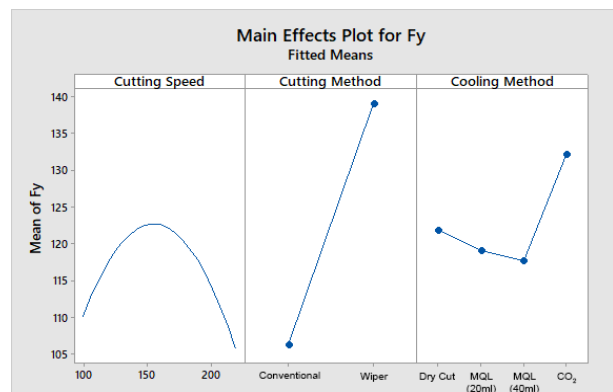


Figure 11. Main effects plot for Fy force

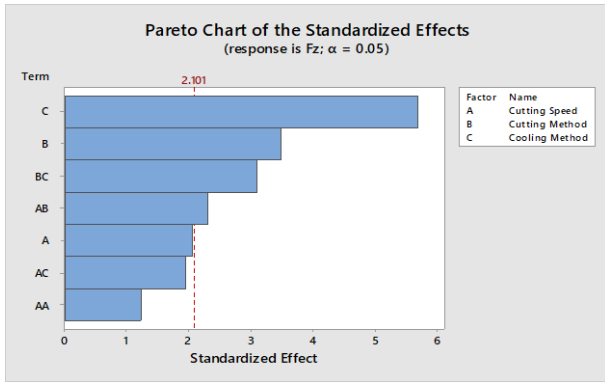


Figure 12. Pareto chart of the standardized effects on Fz force

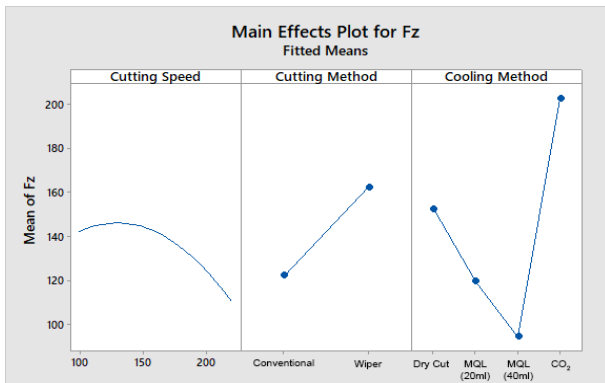


Figure 13. Main effects plot for Fz force

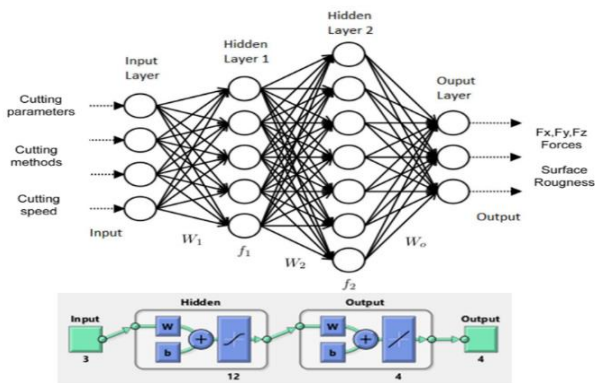


Figure 14. Structure of the trained Artificial Neural Network

To train neurons in the artificial neural network, 70% of the data were reserved for training, 15% of the data were for validation and 15% of the data were for testing. Levenberg-Marquardt method was used for the training of neurons. The Levenberg-Marquardt algorithm is generally used to solve nonlinear least squares problems [25]. Two numerical minimization algorithms are combined into the Levenberg-Marquardt algorithm: the method of gradient descent and the method of Gauss-Newton. The sum of the square errors in the gradient descent method is minimized by updating the parameters in the direction of the steepest descent. The sum of the square errors in the Gauss-Newton method is reduced by assuming that the function of the least squares is locally quadratic in the parameters and finding the minimum of that quadratic. When the

parameters are far from their appropriate value, the Levenberg-Marquardt method behaves more like a gradient-descent technique and acts more like the Gauss-Newton method when the parameters are close to their optimal value. The results obtained are given in Table 2.

The R value for all training, validation and testing procedures was found to be 0.99 and above. R value close to 1 shows that the model performance is very high. The error distribution of the developed artificial neural network model has been quite smooth and is given in Figure 15.

The regression results of the artificial neural network are also given in Figure 16. The fact that the data points are located on the diagonal and not scattered shows the success of the developed model.

Cutting speed and MQL values can be set to a desired value with sliding bar components. When the calculate button is pressed, the values of Ra, Fx, Fy and Fz which the artificial neural network has predicted with high performance can be read on the interface with the input values received from the user. With the developed interface, the input parameters can be changed and the surface roughness and axial forces change as a result of machining can be easily examined.

MATLAB App Designer tool was used to develop an interactive interface for the end users. With this tool, the user can select the cooling type in the cutting parameters section and the processing method in the cutting methods section via radio buttons. The screenshot of the developed interface is given in Figure 17.

Table 2. Performance parameters for the developed ANN

Neural Network Results	MSE	R
Training	24.239	0.998
Validation	36.046	0.996
Testing	46.788	0.997

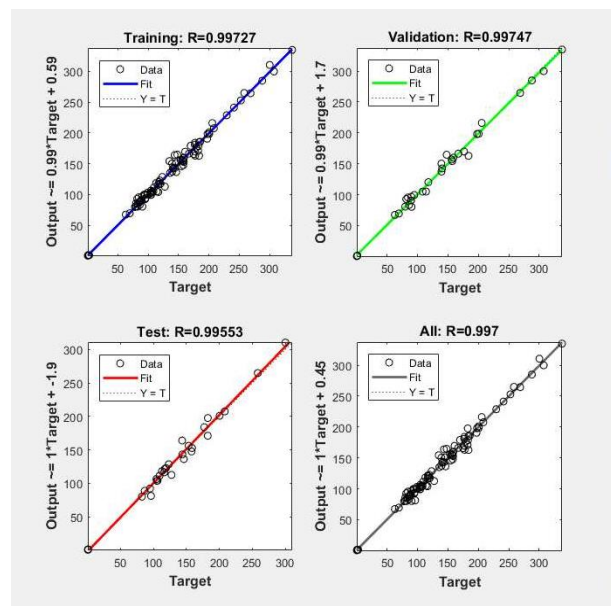


Figure 15. Error histogram for the ANN

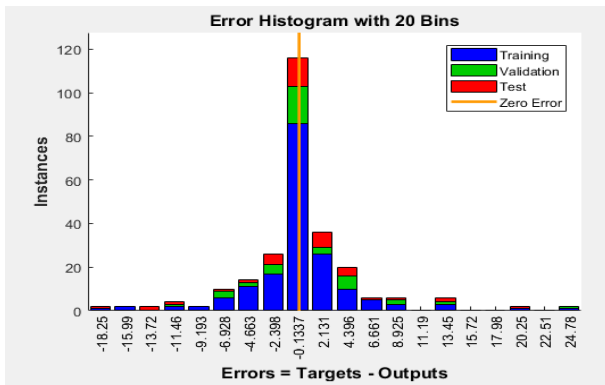


Figure 16. Regression results for the developed model.

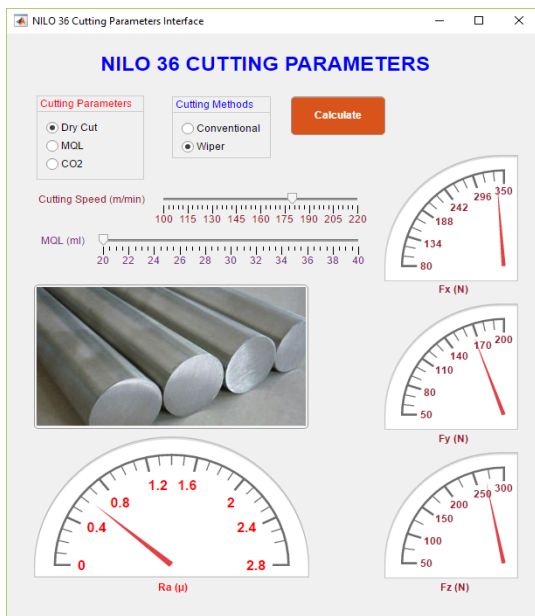


Figure 17. The interactive interface for Nilo 36 cutting parameters

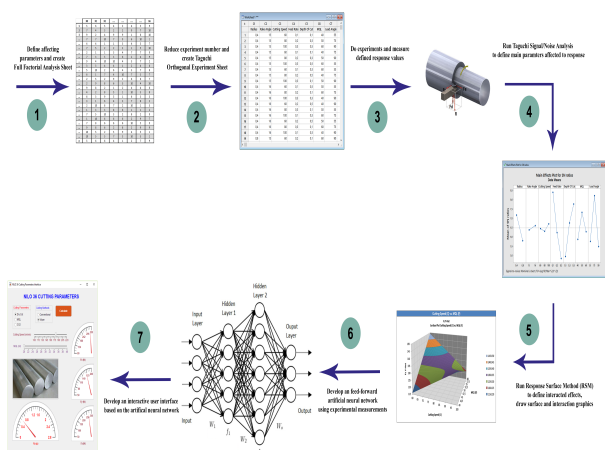


Figure 18. The graphical summary of experimental workings

4. Conclusion

In this study, Nilo 36 superalloy which is widely used in avionics sector has been processed by using 2 different methods (conventional and wiper) at different cutting

speeds between 100 and 220 rpm. Dry cut, MQL (20ml and 40ml) and CO₂ cooling methods were applied during the cutting process and what effects the machining parameters have on the surface and the forces were investigated. A full factorial analysis was performed to plan the parameters to be measured before the experimental measurements were performed. Then, Taguchi method was used to reduce the number of experiments and the required number of experiments was determined as 32.

The results obtained from the measurements were examined by response surface methodology and ANOVA analysis and the effect levels of the factors affecting the results were calculated. The results are presented in graphs. After all these analyses, the results were loaded onto a feed-forward artificial neural network and it has been trained with measurement data. The performance of the trained network was calculated in terms of MSE and R parameters and the R value was found as 0.99.

After the artificial neural network developed, it is connected to an interface where the user can interact and see the interaction results instantly on the screen. The results obtained within the scope of the study, developed artificial neural network model and interface application revealed a unique and useful tool. Thus, the user can examine the effect of all cutting parameters on cutting forces and surface roughness easily.

Declaration

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article. The authors also declared that this article is original, was prepared in accordance with international publication and research ethics, and ethical committee permission or any special permission is not required.

Author Contributions

G. Basmacı and M. Ay both realized the experimental setup and all measurements. İ. Kirbaş also developed the artificial neural network model and the interactive user interface.

Acknowledgment

Experiments were carried out by using the experimental equipment taken Department of Mechanical Engineering within the scope of FEN-E 090517-0273 project supported by BAPKO of Marmara University, Turkey.

Nomenclature

- MQL* : Minimum quantity lubrication
- Ra* : Surface roughness (µm)
- F_x* : Force applied to “x” axis (N)
- F_y* : Force applied to “y” axis (N)

F_z : Force applied to “z” axis (N)
 N : Cutting force (N)
 MSE : Mean squared error
 rpm : Revolutions per minute

References

- Ha, T.K. and Min, S.H., *Effect of C Content on the Microstructure and Physical Properties of Fe-36Ni Invar Alloy*, Eco-Materials Processing and Design XV, Trans Tech Publications Ltd, 2015. p. 293–296.
- Maranhão, C. and Davim, J.P., *Finite Element Modelling of Machining of AISI 316 Steel: Numerical Simulation and Experimental Validation*, Simulation Modelling Practice and Theory, 2010. **18**(2): p. 139–156.
- Tekaslan, Ö., Gerger, N., Günay, M. and Şeker, U., *Examination of the Cutting Forces of AISI 304 Austenitic Stainless Steel in the Turning Process with Titanium Carbide Coated Cutting Tools*, Pamukkale Univiversitesi Muhendislik Bilim. Dergisi, 2007. **13**(2): p. 135–144.
- Li, D.W., Chen, H.T., Xu, M.H. and Zhong, C.M., *Study on Turning Parameter Optimization of Austenitic Stainless Steel*, Mechanical Engineering and Green Manufacturing, Trans Tech Publications Ltd, 2010. p. 1829–1833.
- Diniz, A.E., Ferreira, J.R. and Filho, F.T., *Influence of Refrigeration/Lubrication Condition on SAE 52100 Hardened Steel Turning at Several Cutting Speeds*, International Journal of Machine Tools and Manufacture, 2003. **43**(3): p. 317–326.
- Kaladhar, M., Subbaiah, K. and Rao, Ch.S., *Optimization of Surface Roughness and Tool Flank Wear in Turning of AISI 304 Austenitic Stainless Steel with CVD Coated Tool*, Journal of Engineering Science Technology, 2013. **8**: p 165–176.
- Dirviyam, P.S. and Palanisamy, C., *Optimization of Surface Roughness of AISI 304 Austenitic Stainless Steel in Dry Turning Operation Using Taguchi Design Method*, Journal of Engineering Science Technology 2010. **5**: p 1-9.
- Basmacı, G., Ay, M. and Kırbaş, İ., *Optimisation of Machining Parameters in Turning 17-4 Ph Stainless Steel Using the Grey-Based Taguchi Method*, Erzincan Üniversitesi Fen Bilimleri Enstitüsü Dergisi 2017. **10**(2): p. 243–254.
- Ay, M., *Optimisation of Machining Parameters in Turning AISI 304L Stainless Steel by the Grey-Based Taguchi Method*, Acta Physica Polonica A, 2017. **131**: p. 349–354.
- Basmacı, G., Ay, M., *Optimization of Cutting Parameters, Condition and Geometry in Turning AISI 316L Stainless Steel Using the Grey-Based Taguchi Method*, Acta Physica Polonica A, 2017. **131**: p. 354–359.
- Dhar, N.R., Kamruzzaman, M. and Ahmed, M., *Effect of Minimum Quantity Lubrication (MQL) on Tool Wear and Surface Roughness in Turning AISI-4340 Steel*, Journal of Materials Processing Technology, 2006. **172**(2): p. 299–304.
- Itoigawa, F., Childs, T.H.C., Nakamura, T. and Belluco, W., *Effects and Mechanisms in Minimal Quantity Lubrication Machining of an Aluminum Alloy*, Wear., 2006. **260**(3): p. 339–344.
- Qin, S., Li, Z., Guo, G., An, Q., Chen, M. and Ming, W., *Analysis of Minimum Quantity Lubrication (MQL) for Different Coating Tools during Turning of TC11 Titanium Alloy*, Materials, 2016. **9**(10): p. 1-13.
- Sampaio, M.A., Machado, Á.R., Laurindo, C.A.H., Torres, R.D. and Amorim, F.L., *Influence of Minimum Quantity of Lubrication (MQL) When Turning Hardened SAE 1045 Steel: A Comparison with Dry Machining*, International Journal of Advanced Manufacturing Technology, 2018. **98**(1):p. 959–968.
- Basmacı, G., Kurt, M., Ay, M. and Bakir, B., *Optimization of the Effects of Machining Parameters in Turning on Hastelloy C22 Composition through Taguchi Response Surface Methodology*, Acta Physica Polonica A, 2018. **134**, p. 28–31.
- Mostafa, Y., Elbestawi, M.A., Veldhuis, S. C., *Density and mechanical properties in selective laser melting of Invar 36 and stainless steel 316L*, Journal of Materials Processing Technology, 2019. **266**: p. 397-420.
- Kırbaş, İ., Peker, M., Basmacı, G. and Ay, M., *Predictive Modeling and Optimization of Cutting Forces Through RSM and Taguchi Techniques in the Turning of ASTM B574 (Hastelloy C-22)*, Handbook of Research on Predictive Modeling and Optimization Methods in Science and Engineering, IGI Global, 2018. p. 398–417.
- Asiltürk, İ., Neşeli, S. and İnce, M.A., *Optimisation of Parameters Affecting Surface Roughness of Co28Cr6Mo Medical Material during CNC Lathe Machining by Using the Taguchi and RSM Methods*, Measurement, 2016. **78**:p. 120–128.
- Abou-El-Hossein, K.A., Kadirgama, K., Hamdi, M. and Benyounis, K.Y., *Prediction of Cutting Force in End-Milling Operation of Modified AISI P20 Tool Steel*, Journal of Materials Processing Technology, 2007. **182**(1–3): p. 241–247.
- Navneet, K., Gandhi, A., Nakum, B., Anil, S., *Optimization And Analysis of Surface Roughness for Invar-36 End Milling Operations*, Materials Today, 2018. **5**: p. 5281-5288.
- Dahbi, S., Ezzine, L., Moussami, H.E., *Modeling of cutting performances in turning process using artificial neural networks*, International Journal of Engineering Business Management, 2017. **9**: p. 184-196.
- Çakiroğlu, R., Yağmur, S., Acir, A., Şeker, U., *Modelling of Drill Bit Temperature and Cutting Force in Drilling Process Using Artificial Neural Networks*, 2017. **20**(2): p. 333-340.
- Kılıç, F., *Effects of three drying methods on kinetics and energy consumption of carrot drying process and modeling with artificial neural networks*, Energy Sources, Part A: Recovery, Utilization, and Environmental Effects, 2020. p. 1-18.
- Filippis, L.A.C.D., Serio, L.M., Facchini, F., Mummolo, G., *ANN Modelling to Optimize Manufacturing Process*, InTech, 2018. p. 32-38.
- Gavin, H.P., *The Levenberg-Marquardt algorithm for nonlinear least squares curve-fitting problems*, Lecture Notes, Duke University, 2020. <http://people.duke.edu/~hpgav/in/ce281/lm.pdf>.



e-ISSN: 2618-575X

INTERNATIONAL ADVANCED RESEARCHES
and
ENGINEERING JOURNALJournal homepage: www.dergipark.org.tr/en/pub/iarejInternational
Open Access Volume 05
Issue 01

April, 2021

Research Article

Analysis and balancing of assembly line in a machine molding factory**Esra Can** ^{a,*} and **Adalet Öner** ^b ^aDepartment of Industrial Engineering, Bursa Uludag University, Bursa, 16059, Turkey^bDepartment of Industrial Engineering, İzmir Yasar University, İzmir, 35030, Turkey

ARTICLE INFO

Article history:

Received 26 July 2020

Revised 14 October 2020

Accepted 14 November 2020

Keywords:

Assembly line balancing

Mathematical modelling

Ranked positional weight
method

Time study

U-type line balancing problem

ABSTRACT

In industrialization, to be able to make cheap and fast production, assembly lines are one of the most basic elements in serial production systems. It is important to balance the assembly line to continue production smoothly. By assembly line balancing is created, each work step is grouped, stations are created and each station time is brought close to the station cycle times. In this study, a refrigerator top panel pressing line is analysed. The study's aim is balancing the line for increase production rate. Firstly, the line is observed and some studies are planned. A time study is done to analyse the current situation of the line. Time study data are calculated by using Excel. Ranked Positional Weight Method is used as an intuitive method for single model U type assembly line balancing problem and mathematical modelling method is applied. The methods are used to balance the line using time study data. The solution of mathematical modelling is obtained by using Lingo. Results are compared and they are observed that results have almost the same. In conclusion, an assembly line balancing problem is mentioned in this study. Various programs related to the applied methods were used, and the data obtained as a result of current and final calculations were compared. First and last calculations and results are verified with each other. It was seen that the data obtained as a result of the study provided improvement.

© 2021, Advanced Researches and Engineering Journal (IAREJ) and the Author(s).

1. Introduction

Technological advances and the global trade in recent years, strict price and quality competition, concentrated on production methods entail a number of new functions. Today, with the impact of rapidly increasing world population level, life is getting much more complex with the requirement to produce a large number of products. Therefore, in the middle of the 20th century "assembly line balancing" idea was posed that the amount of the claim as soon as possible, as a result of the effort to produce the desired quality and cheap. Companies must make the best use of their assembly lines in today's competitive environment. The importance of assembly line balancing also occurs here. Assembly line workstations are systems that combine a material handling system. The purpose of the system is to assemble the components of a product and obtain the finished product [1]. When the line of the product to be assembled is designed, the problem of

balancing the time differences arises between the operations of the product. Assigning work elements to stations for this purpose is called 'assembly line balancing problem'. During the creation of the product need to be done jobs, should be assigned to the assembly stations to minimize assembly line balancing times. The installation of a work element should be done by looking at the priority relationship at the predetermined station [2].

Various classifications can be made as single, multiple or mixed models depending on the number of models produced on the assembly line. Assembly line balancing problems can be classified into two groups according to the status of duty times; stochastic and deterministic assembly lines. When an assembly line is fully automated, all the tasks will have a fixed operation time. When tasks are performed manually at the workstations, variability (or stochasticity) emerges [3]. There can be two main goals while balancing an assembly line. The goals are

* Corresponding author. Tel.: +90 224 294 2081.

E-mail addresses: esracan545@gmail.com (E. Can), adalet.oner@yasar.edu.tr (A. Öner)

ORCID: 0000-0002-4150-346X (E. Can), 0000-0001-5989-3825 (A. Öner),

DOI: 10.35860/iarej.772678

This article is licensed under the CC BY-NC 4.0 International License (<https://creativecommons.org/licenses/by-nc/4.0/>).

minimisation of the number of workstations for given cycle time and minimisation of the cycle time for a given number of workstations [4].

As already mentioned above, types of assembly line balancing are based on a set of limiting assumptions. Some of those assumptions; the processing sequence of tasks is subject to precedence restrictions, no assignment restrictions of tasks besides precedence constraints, all stations are equally equipped for machines and workers [5].

The manufacturing / installation line balancing problem is classified in different types according to the shape of the line. Assembly lines can be designed in different forms such as straight, circular, random, different angle, U-typed, zigzag according to their physical location. Conventional assembly lines are designed straight. Later, the U-typed line was preferred more in new production lines. Assembly lines are placed in u shaped lines in u shaped assembly line balancing problem. The entrance and the exit of the line are on the same parallel position in a U-type line design [6]. In recent years, the just-in-time production philosophy has been spread, U-type assembly lines usage rate has been increased [5].

The assembly line balancing method should be inspected in two groups in terms of solution method; analytical methods and heuristic methods. Analytical methods are known as optimization methods. The methods comprise of objective function and constraints. Heuristic methods give an approximate solution to the problems. There are some heuristic methods in literature such as the Ranked Positional Weight Method (Helgeson-Birnie), Precedence relationship diagram (Hoffman), COMSOAL technique (Arcus), Two-phase balancing technique (Moddie-Young), Probabilistic line balancing (Elsayed-Bouch), Grouping technique (Tonge) [7].

Assembly line balancing is mostly used the production systems. The first article about assembly line balancing was written by Salveson. In this study 0-1 integer programming model was improved to solve the problem [8]. Alagas et al studied a new constraint programming model for mixed-model assembly line balancing problems. The model minimizes the cycle time for a given number of stations [9]. Nicosio et. al. studied the problem of assigning operations to an ordered sequence of non-identical workstations, which also took precedence relationships and cycle time restrictions into consideration. The study aimed to minimise the cost of workstations. They used a dynamic programming algorithm and introduced several rules to reduce the number of states in the dynamic program [10]. Helgeson and Birnie developed a heuristic method that is called the 'Ranked Positional Weight (RPW) Method' in 1961. This method provides to assign work elements to the station in an optimal way. It takes into account the precedence relationships as well as

the processing time of all tasks. The RPW value of each operation is determined and assigned operations to workstations in this method [11]. In the literature, besides the RPW method, heuristic algorithms such as tabu search are also used in line balancing problems [12]. However, when looking at the studies done, one of the most commonly used heuristic line balancing methods in the industry is the RPW approach [13]. Bongomin et al used the RPW method at balancing a trouser assembly line to increase the line efficiency as well as minimize the number of workstations. According to scenario 1, after calculating the existing assembly line the results show that the line efficiency is 35.66%, the last situation gave an improvement of line efficiency which is 80.56% by using the RPW method. The balance delay reduces from 64.34 to 19.44 and workstation number decreases from 61 to 27 in the study [14]. Ikhsan used the RPW method to develop and balance the assembly line. The number of actual workstations is 9 and it decreases to 8 in the last situation. Idle time reduce from 10441 to 3286. While Line efficiency and balance delay are 66.08%, 33.92% respectively in the initial situation, they become 86.09%, 13.91% after the RPW calculations [15]. The studies show that minimization of workstations contributes to an increase in line efficiency. In addition to the RPW method, we tackled the problem with mathematical modelling and supported the result with two approaches in the study. While dealing with the problem in the study, in addition to the RPW, which is one of the solution methods in the literature, the solution was supported by an analytical method by using a mathematical modelling method. Since there are not many studies in which these two methods are used together, it will offer a different perspective for researchers.

In this study, the role and usage method of assembly lines in the production system is mentioned. The problem of assembly line balancing is addressed frequently encountered in production lines in factories. A problem is defined related to the factory's order in a line where is the pressing process of refrigerator top panel in the molding factory. The problem arises from idle time, bottleneck between the stations and unbalanced line flow. As a result of these problems, the order is not satisfied on time and firms wait for their orders for a long time. This study was deemed necessary to eliminate these problems arising on the lines and to meet customer orders on time. It is aimed to solve the problem by using two techniques that are the "RPW" and "mathematical modelling" while balancing the assembly line. In the last part, the results have been evaluated and mentioned the effect on real life.

2. Material and Method

The assembly line balancing problem was handled with analytical and heuristic methods in this study. Heuristic

methods like the RPW method present close solutions to the problems. Balancing the assembly line two techniques were used which are the "RPW method" and "mathematical modelling". The modelled problem was solved by the time study's data. In the study, the RPW method approach was supported by a mathematical modelling approach, and solutions were compared.

2.1 Time Study Calculations

Time study is a method for recording times and rates of working for specific job elements to build standards. The time study aims to establish a time for a qualified worker to perform specified work under stated conditions and at a defined rate of working [16]. There are 34 work elements in the stations. The works' times were observed by using a stopwatch for 20 observations. Observation time, normal time, average observed and normal time, standard time, and standard deviation calculated for each element. All of these calculations are shown in time study forms in appendix (Table A.1).

The observation number was calculated at a 95 percent confidence level with acceptable error 0.05 (k). While analyzing the sample size of observation, t distribution table is used [14]. The observation number is calculated by this formula. s is standard deviation; t is distribution value; k is an acceptable fraction for error; \bar{x} is the mean of the sample observations.

$$n' = \frac{S * t}{k * x} \tag{1}$$

The required observation numbers' notations and formula are shown below.

N=sufficient number of observations

n'=number of observation taken

x_i = time measured in i.th observation.

40 (hours) = 8(shift)*5(day)

$$N = \left[\frac{40 \sqrt{n' \sum x_i^2 - (\sum x_i)^2}}{\sum x_i} \right]^2 \tag{2}$$

If the required observation number is higher than the observed number, the observation number should be tested once more time. The observation time values calculation is shown in a table in Appendix (Table A.1).

2.1.1 Performance Rating Calculations

The procedure is for determining the value for a factor which will adjust the measured time for an observed task performance to a task time that one would expect of a trained operator performing the task, utilizing the approved method and performing at a normal pace under specified workplace conditions [17]. Performance rating

shows, level of workers' motivation. The most convenient method to designate performance rating is the Westinghouse System. According to Westinghouse System Method, there are four factors in evaluating the performance of the operator. These are skill, effort, conditions, and consistency [16, 17]. The system factors' ratings are shown in the performance rating in Table 1.

If the method is applied for the first work element;

Rating Factor = 1 \mp Westinghouse Rating,

According to these data performance rating is;

$$100+100*(0.02-0.04+0.00-0.1) = 88.$$

The performance rating values are calculated with this formula for each operator.

2.1.2 Calculation of the Normal Time

Normal time is calculated by this formula:

$$Normal\ Time = Observed\ time * \frac{Performance\ rating}{100} \tag{3}$$

If the first work element's normal time is calculated by this formula;

$$1^{st}\ element's\ normal\ time\ is\ calculated\ as\ 1.5*(88/100) = 1.32\ second.$$

2.1.3 Calculations of the Allowances

There are various allowances to build standard time. These are personal, fatigue, delay allowances. Calculations of the allowances are shown below.

Personal Allowances Calculations:

$$Normal\ Time = Observed\ time * \frac{Performance\ rating}{100} \tag{4}$$

A daily work hour is a total of 420 minutes in the factory. Calculation of Personal Allowances are given in Table 2.

Table 1. Performance ratings for the first element [16].

Effort	E1	Fair	- 0.04
Conditions	C	Good	+0.02
Consistency	D	Average	0.00
Skill	E2	E2	-0.1
Total Rating Factor			-0.12

Table 2. Calculations of personal allowances

Personal Allowances	Times in a day
Taking a drink	7 minutes
Going to bathroom	8 minutes
Talking to other workers	7 minutes
Total Personal Allowances	22 minutes
Personal allowances percentage [(22 minutes/420 minutes)*100]	5.23%

Fatigue Allowances Calculations:

Fatigue Allowances

$$Percentage = \frac{Fatigue\ allowances\ time}{Daily\ work\ minute} * 100 \tag{5}$$

Calculation of Fatigue Allowances are given in Table 3.

Delay Allowances Calculations:

Delay Allowances Calculations = (Delay Allowances Time/ Daily Work Minute)*100

Total allowances=Personal allowances+ Fatigue Allowances+ Delay Allowances.

Calculation of Fatigue Allowances are given in Table 4.

2.1.4 Calculations of the Standard Times

While the standard time is calculated; assigned to a particular employee is a qualified job, tools and equipment, raw materials, the work environment, and the level of performance are needed to be defined. Standard Time is calculated using the following formula.

$$Standard\ Time = Normal\ Time * (1 + Allowances)$$

For example for the first element performance rating is below;

$$Standard\ Time = 1.52 * (1 + 0.16) = 1.766 \cong 1.77\ seconds.$$

Other work elements' standard time is calculated in the same way as the first work element. The total standard times are calculated as 100.22 seconds.

For obtaining and saving data time study and the stopwatch methods are used. In the time study form model name work element name and number, operator name machine, machine number, performance ratings are written. For every work element in the assembly line, 20 observations were made and written on to the time study form (Table A.2). A time Study form was prepared for work element 1. The time study form of every task is shown in Appendix (Table A.1).

Table 3. Calculations of fatigue allowances

Fatigue Allowances	Times in a day
Work breaks (tea or coffee breaks)	30 minutes
Total Fatigue Allowances	30 minutes
Fatigue Allowances percentage [(30 minutes/420 minutes)*100]	7.14%

Table 4. Calculations of delay allowances

Delay Allowances	Times in a day
Machine injury	5 minutes
Machine maintenance and breakdown	6 minutes
Waiting for materials	3 minutes
Controlling and cleaning work place	2 minutes
Total Delay Allowances	16 minutes
Delay Allowances (16 minutes/420 minutes)*100	3.80%
Total allowances=5.23%+7.14%+3.80% (Personal+Fatigue+Delay Allowances)	16.17%

2.2 Application of Assembly Line Balancing

The work study processes are applied for line balancing. The standard times are calculated by using the work study method. According to observations which are done in the factory determined operator's number, machine numbers, and work hours in the assembly line in this factory. The factory works from 8:45 to 16:45. There is an assembly line that has 34 operations, 6 machines, and 7 operators in the production area. These machines are 2 excentric and 4 hydraulic presses. Sixth and seventh operations are done in the same machine because the machine has two processes sides.

The factory wants to increase the production rate. An assembly line balancing study was done in the factory to meet demand. Firstly, some calculations were made, such as calculating cycle times, drawing the priority relationship diagram, and applying the sequential positional weight method. Then, the mathematical modelling method was used and compared the two methods' results.

2.2.1 Ranked Positional Weight Method

The RPW method takes account of the standard times value of the element and its position in the precedence diagram. Then the elements are assigned to the workstation in the general order of their RPW values. Precedence Relationship Diagram and standard time of work elements are calculated and shown in appendix (Table A.3). The precedence matrix of the top panel is shown in appendix (Figure A.1).

2.2.2 Calculation of Cycle Time

The company starts to work from 8:45 am to 16:45. A shift period takes 8 hours. There are a one-hour lunch break and 2 coffee/tea breaks. Every coffee or tea break takes 15 minutes. According to these data, available working time and cycle time are calculated. "C" states cycle time, "N" is production rate, and "T" is daily work hours (Available working time).

$$C = \frac{T}{N} \tag{6}$$

N=1150;

Available working time= (8 hours*60 minutes)-(30 minutes +15 minutes +15 minutes) = 480minutes-60 minutes = 420 minutes = 25200 seconds.

C = 25200/1150 = 21.90 seconds. After calculation of cycle time, workstation number has been calculated.

$$Workstation\ number = WS = \frac{(Total\ Production\ Time)}{(Cycle\ Time)}$$

$$WS = 100.22 / 21.90 = 4.57 \cong 5\ workstations$$

Calculations of "RPW" and "distribution of the tasks" are shown in Table 5.

Table 5. Calculation of RPW and distribution of tasks

Station No	Tasks	Cumulative Station Time	Weight Distribution	Ranked Positional Weight
1st station	1	1.76	1.76	100.22
	2	3.64	5.40	98.34
	3	0.95	6.35	94.70
	4	0.79	7.14	93.75
	5	2.38	9.52	92.96
	6	3.42	12.94	90.58
	7	3.98	16.92	87.16
	8	1.04	17.96	83.18
	9	3.34	21.30	82.14
2nd station	10	3.66	3.66	78.80
	11	2.28	5.94	75.14
	12	2.18	8.12	72.86
	13	0.68	8.80	70.68
	14	9.31	18.11	70.00
	15	1.99	20.10	60.69
3rd station	16	2.23	2.23	58.70
	17	3.17	5.40	56.47
	18	1.59	6.99	53.30
	19	0.57	7.56	51.71
	20	4.13	11.69	51.14
	21	2.03	13.72	47.01
	22	5.13	18.85	44.98
	23	1.78	20.63	39.85
4th station	24	5.04	5.04	38.07
	25	4.07	9.11	33.03
	26	2.30	11.41	28.96
	27	2.05	13.46	26.66
	28	2.17	15.63	24.61
	29	3.59	19.22	22.44
5th station	30	4.43	4.43	18.85
	31	1.61	6.04	14.42
	32	5.12	11.16	12.81
	33	3.03	14.19	7.69
	34	4.66	18.85	4.66

The calculation of idle time is as follows. “C” states cycle time, “D” is the difference between task and cycle time, “T” is workstation time (task time), “K” is station number, and “d_k” is the delay time of kth station.

$$d_k = C - Tk \quad (7)$$

$$D = k_c - \sum ti \quad (8)$$

Total Idle Time = $\sum D = (21.9-21.30) + (21.9-20.10) + (21.9-20.63) + (21.9-19.22) + (21.9-18.85)$

$$\sum D = 9.40 \text{ seconds.}$$

Proportion of idle time for each cycle = $9.40/(21.90*5) = 0.085 = 8.5\%$.

2.2.3 Calculation of Balance Efficiency and Balance Delay

There are some notations for the calculations of balance efficiency and delay as follows. D is balance delay; t is time of the work elements; E is efficiency; C is cycle time; K is number of work station.

Balance Efficiency;

$$E = [(\sum ti)/(K * C)]*100 \quad (9)$$

Balance Delay;

$$D = [(K * C - \sum ti)/(K * C)]*100 \quad (10)$$

Before the line balancing:

$$\text{Balance Efficiency; } E = [100.2/(7 * 21.90)]*100=0.65$$

$$\text{Balance Delay; } D = [(7 * 21.90 - 100.22)/(7 * 21.90)] *100=0.3462$$

After the line balancing:

$$\text{New Balance Efficiency; } E = [100.22/(5 * 21.90)]*100=0.915$$

$$\text{New Balance Delay; } D = [(5 * 21.90 - 100.22)/(5 * 21.90)]*100=0.0847$$

2.3 Mathematical Modelling Method

The assembly line balancing is an optimization problem. The proposed mathematical model is solved with the Lingo program providing solution near to optimal solutions. “Task” defines work elements, “Station” defines station number, and “Predecessor” defines the precedence relationship between works. While writing the constraints, have been taken into account the criteria below. The mathematical model is formulated for our problem.

- The sum of task times for the set of tasks assigned to each workstation does not exceed the cycle time (Equation 11) [18].
- Each task must be assigned to only one workstation (Equation 12) [18].
- The stations must be arranged according to a precedence relationship. If task v is to be assigned to station b, then it is immediate predecessor u must be assigned to some station between 1 and b (Equation 13) [18]. These constraints are among the strict constraints in the model.

In this model, the station number is assigned to a big number “M (100.000)” (Equation 14). The set of P identifies the assembly order constraints.

There are some notations in the model:

P = {(u,v): task u must precede task v}.

K: Number of workstation

N: Number of tasks

R: Required number of workstation.

C: Cycle time

The notations considered in the mathematical model are clarified as follows:

Indices

i, u, v: tasks

k, b, j: stations

t: Time of the work elements

t_i: Completion time of the task "i"

Decision Variables

$$x_{ik} = \begin{cases} 1; & \text{if task } i \text{ is assigned to station } k \\ 0; & \text{otherwise;} \end{cases}$$

Constraints

$$\sum_{i=1}^N t_i X_{ik} \leq C \quad k = 1, \dots, K \quad (11)$$

$$\sum_{k=1}^K X_{ik} = 1 \quad i = 1, \dots, N \quad (12)$$

$$X_{vb} \leq \sum_{j=1}^b X_{uj} \quad b = 1, \dots, K \text{ and } (u, v) \in P \quad (13)$$

$$\sum_{i=1}^N X_{ij} \leq MR_j \quad \forall_j \quad (14)$$

$$x_{ik} \in \{0,1\} \quad \forall_i; \forall_k \quad (15)$$

The optimization problem is formulated using binary variables as decision variables. Equation (15) denotes the constraint showing that the decision variable consists of 0-1 integer variables. The domain of the variables is defined in this way.

Objective Function

$$\text{Min} \sum_{j=1}^K R_j \quad (16)$$

3. Results and Discussion

Some data were obtained as a result of the calculations. The mathematical model and RPW method’s results were calculated and comprised by using the data. According to the RPW method’s results, when the initial situation and the last situation are compared it is clear that while the workstation number decreases, the line efficiency increases. Total idle time had been calculated 53.08 seconds in the actual situation. After balancing, the total idle time had decreased to 9.40 seconds. Balance efficiency was improved from 65% to 91.52%, and balance delay was decreased from 34.62% to 8.47%. It shows that, when the workstation numbers drop from 7 to 5, time losses decrease significantly. The values of the initial and the last situation of parameters are shown in Table 6 comparatively.

The objective of the mathematical model is to minimize the number of stations. The model is applied for two cycle time “22” and “35”. Additionally, to see the effect of the increase in cycle time on the number of stations, the model was also solved with a high cycle time such as 35. The distribution of tasks to each station is shown in the two situations clearly. The mathematical model is calculated for the cycle time 22 and then for cycle time 35.

Table 6. Comparison of the parameters before and after balancing

Parameters	Initial Situation	Last Situation
Number of workstations	7	5
Total idle time (sec.)	53.08	9.40
Balance delay (%)	34.62%	8.47%
Balance efficiency (%)	65%	91.52%

Table 7. Results of mathematical model on lingo

Global optimal solution for "cycle time 22"	Global optimal solution for "cycle time 35"
Objective value: 5.000000	Objective value: 4.000000
Objective bound: 5.000000	Objective bound: 4.000000
Infeasibilities: 0.000000	Infeasibilities: 0.000000
Extended solver steps: 129	Extended solver steps: 267
Total solver iterations: 15883	Total solver iterations: 41316

According to lingo solutions, the objective function value is 5 for the cycle time 22. If the cycle time is increased to 35, the station number is to be 4. It means that, if the cycle time is increased, the station number is decreased. These calculations' results of lingo solutions are shown in Table 7. However, "cycle time 22" and "5 stations" are more optimal to balance both cycle times and the number of stations.

Similarly, with the RPW method solution, the station number is decreased from 7 to 5 stations in the mathematical model with cycle time 22. With the decrease in the number of stations, line efficiency increased and delays decreased. In this way, the mathematical modelling solution is near and parallel to the RPW method solution.

4. Conclusion

In this project, an assembly line was analysed. At the end of the literature review and researches, some methods were determined to solve the assembly line balancing problem. The RPW method and the mathematical modelling methods were used. These methods were compared and it was found that the results similar to each other. As a result of the studies, after line balancing, balance efficiency was improved from 65% to 91.52%. Balance delay decreased from 34.62% to 8.47% according to the current situation of the line. At the same time, idle times were decreased. In the mathematical modelling solution, the station number decreased from 7 to 5 stations similar to the RPW method solution. The mathematical modelling method was applied with two different cycle times to see the difference between changes in station numbers. While station number decreases, idle times decrease. The order will be satisfied as requested. It has been found that the proposed solution provides a significant improvement in assembly line efficiency. It is thought that this study will provide a different perspective for researchers dealing with assembly line balancing problem. In future studies, in addition to minimizing the number of stations, new constraints will be added to the model and effective and efficient use of stations will be provided. Moreover, in future studies, in addition to mathematical models and RPW solutions, artificial intelligence algorithms will be used to achieve more optimal results faster in the solution space.

Declaration

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article. The authors also declared that this article is original, was prepared in accordance with international publication and research ethics, and ethical committee permission or any special permission is not required.

Author Contributions

E. Can developed the methodology, collected data, and formed the inputs of the mathematical model. She worked in the time study process, applied heuristic algorithms and analyzed the data. A. Öner structured the mathematical model and contributed to other processes.

References

- Li, M., Tang, Q., Zheng, Q., Xia, X., Floudas, C. A. *Rules-based Heuristic Approach for the U-shaped Assembly Line Balancing Problem*. Applied Mathematical Modelling, 2017. **48**: p. 423-439.
- Kilinci, O. *Petri net-based algorithm for maximizing production rate in assembly lines*. Journal of the Faculty of Engineering and Architecture of Gazi University, 2020. **35**(2): p. 753-763.
- Alagas, H. M., Yuzukirmizi, M., Turker, K., T. *Stokastik Montaj Hatlarinin Kisit Programlama ve Kapali Kuyruk Aglari ile Dengelenmesi*. Journal of the Faculty of Engineering Architecture of Gazi University, 2013. **28**(2): p. 231-240.
- Fathi, M., Fontes, D. B. M. M., Moris, M. U., Ghobakhloo, M. *Assembly Line Balancing Problems: A Comparative Evaluation of Heuristics and a Computational Assesment of Objectives*. Journal of Modelling in Management, 2018. Available from: <https://doi.org/10.1108/JM2-03-2017-0027>.
- Boysen, N., Fliedner, M., Scholl, A. *A Classification of Assembly Line Balancing Problems*. European Journal of Operational Research, 2007. **183**: p. 674-693.
- Delice, Y., Aydogan, E., Ozcan, U. *Stochastic Two-Sided U-Type Assembly Line Balancing: A Genetic Algorithm Approach*. International Journal of Production Research, 2016. **54**(11): p. 3429-3451.
- Kayar, M., Akyalcin, Ö. C. *Applying Different Heuristic Assembly Line Balancing Methods in the Apparel Industry and their Comparison*. Fibres & Textiles in Eastern Europe, 2014. **6**(108): p. 8-19.
- Salveson, M.E. *The Assembly Line Balancing Problem*. Journal of Industrial Engineering, 1955. **6**(3): p. 18-25.
- Alagas, H. M., Pinarbasi, M., Yuzukirmizi, M., Toklu, B. *Karma Modelli Tip-2 Montaj Hattu Dengeleme Problemi İçin Bir Kısıt Programlama Modeli*. Pamukkale University Journal of Engineering Science, 2016. **22**(4): p.340-348.
- Nicosia, G., Pacciarell, D. Pacifici, A. *Optimally balancing assembly lines with different workstations*. Discrete Applied Mathematics, 2002. **118**(2002): p. 99-113.
- Helgeson, W. P., Birnie, D. P. *Assembly Line Balancing Using the Ranked Positional Weight Technique*. Journal of Industrial Engineering, 1961. **12**(6): p. 384-398.
- Durmaz, H., Koyuncu, M. *Optimization of Assignment Problems in Production Lines with Different Skilled Labor*

Levels. International Advanced Researches and Engineering Journal, 2019. 03(02): p. 123-136.

13. Islam, S., Sarker, S., Parvez, M. *Production Efficiency Improvement by Using Tecnomatix Simulation Software and RPWM Line Balancing Technique: A Case Study*. American Journal of Industrial and Business Management, 2019. 9: p. 809-820. Available from: <https://doi.org/10.4236/ajibm.2019.94054>.
14. Bongomin, O., Mwasiagi, J. I., Nganyi, E. O., Nibikora, I. *Improvement of garment assembly line efficiency using line balancing technique*. Engineering Reports, 2020. 2(5): p. 1-22. Available from: <https://doi.org/10.1002/eng2/12157>

15. Ikhsan, S., *Application of ranked positional weights method in springbed production line balancing*. IOP Conf. Series: Materials Science and Engineering, 2020. 801. Available from: <http://doi:10.1088/1757-899X/801/1/012098>.
16. Niebel, W., Freivalds, A. *'Time Study', 'Methods Standard and Work Design'*. 2013, New York: The McGraw-Hill Companies, Inc. p. 415-449.
17. Salvendy, G. *Handbook of Industrial Engineering, Technology and Operations Management*. 3rd. edition. 2001, USA: John Wiley & Sons, p. 2832.
18. Askin, R. G., Standridge, C. H. *'Assembly Lines: Reliable Serial System', 'Modelling and Analysis of Manufacturing Systems'*. 1993, USA: John Wiley & Sons, p. 480.

Appendix

Table A.1. General time study form

Station-1 Burr Cutting Process		OBSERVED TIMES																										
work element no	work element	t value	Standard deviation	Mean	k	n'	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20		
2.09	1 Taking panel from table with magnet	0.372	1.728	0.05	81.1	1.5	1.800	1.320	1.260	1.980	1.900	2.040	1.560	1.620	2.340	1.320	2.400	1.200	1.380	2.340	1.920	1.500	1.980	1.860	1.600	1.740	2.00	
2.09	2 Putting panel into the press	0.240	2.831	0.05	12.5	2.82	2.75	2.33	2.56	2.89	2.71	3.01	2.91	2.36	2.87	3.12	3.24	2.48	2.93	3.06	3.09	2.99	3.06	2.99	2.73	2.79	2.87	2.09
2.09	3 Standing on press button	0.211	0.745	0.05	140	0.66	0.72	0.61	0.92	0.47	0.95	0.81	0.65	0.77	0.68	0.85	0.42	0.47	1.32	1.03	0.57	0.68	0.82	0.68	0.81	0.81	2.09	
2.09	4 Working the press	0.007	0.572	0.05	0.3	0.57	0.57	0.58	0.57	0.58	0.57	0.57	0.57	0.58	0.57	0.57	0.57	0.58	0.56	0.56	0.57	0.59	0.56	0.57	0.57	0.57	2.09	
2.09	Taking the panel from press	0.327	1.954	0.05	48.8	1.32	1.76	1.48	1.52	1.99	1.73	2.06	1.73	2.06	2.14	1.96	2.45	2.06	2.14	1.85	1.67	1.96	2.15	2.26	1.82	1.97	2.13	2.09
2.09	5 stall and putting to the table																										2.09	
Station-2 Border Bending Processes		OBSERVED TIMES																										
work element no	work element	t value	Standard deviation	Mean	k	n'	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20		
2.09	1 Taking panel from table with magnet	0.542	2.949	0.05	59	2.100	2.800	2.63	2.87	2.61	3.21	3.61	2.99	3.01	3.67	2.89	3.27	3.54	4.01	2.21	2.38	3.64	2.54	2.71	2.29	2.29	2.09	
2.09	2 Putting panel into the press	0.407	3.865	0.05	27.2	4.08	3.71	2.96	3.19	3.65	2.94	3.17	2.89	2.86	3.76	3.49	2.81	4.03	2.91	3.25	2.88	3.25	3.54	3.54	3.54	3.54	2.09	
2.09	3 Standing on press button	0.082	0.721	0.05	22.4	0.66	0.68	0.71	0.69	0.77	0.65	0.81	0.91	0.67	0.68	0.74	0.61	0.73	0.68	0.66	0.66	0.67	0.81	0.67	0.90	0.81	2.09	
2.09	4 Working the press	0.420	3.022	0.05	336.3	2.088	2.088	2.088	2.087	2.089	2.088	2.084	2.084	2.084	2.088	2.09	2.091	2.092	2.088	2.088	2.088	2.088	2.089	2.089	2.089	2.089	2.09	
2.09	Taking the panel from press	0.658	2.865	0.05	92.1	2.28	2.54	3.12	3.89	2.75	1.53	1.87	3.65	3.54	3.42	3.22	2.96	2.74	2.51	4.04	2.25	2.31	2.87	3.17	2.63	2.63	2.09	
2.09	5 stall and putting to the table																										2.09	
Station-3 Bending Process		OBSERVED TIMES																										
work element no	work element	t value	Standard deviation	Mean	k	n'	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20		
2.09	1 Taking panel from table	0.583	2.2355	0.05	119	1.44	1.61	1.29	1.84	2.11	2.15	2.64	1.58	1.67	2.66	3.16	2.70	2.41	2.77	2.95	3.05	2.67	1.85	1.69	2.44	2.44	2.09	
2.09	2 Putting panel into the press	0.854	2.086	0.05	293	1.26	1.02	0.94	1.56	1.83	1.33	1.47	2.16	2.83	2.67	1.98	1.65	2.74	3.26	2.67	3.14	2.61	3.65	1.53	1.42	1.42	2.09	
2.09	3 Standing on press button	0.145	0.5313	0.05	130	0.36	0.496	0.65	0.49	0.89	0.52	0.38	0.38	0.46	0.71	0.38	0.46	0.81	0.51	0.37	0.6	0.43	0.55	0.39	0.42	2.09		
2.09	4 Working the press	0.023	8.0156	0.05	0.01	8.02	8.01	8.02	8.04	7.99	8.02	8	8.05	7.98	8.04	8.05	7.99	7.98	8.01	8.03	8.04	7.99	8	8.01	8.01	8.01	2.09	
2.09	Taking the panel from press	1.171	1.9495	0.05	632	0.48	1.32	0.98	0.84	2.65	4.27	3.82	3.64	2.17	0.88	0.79	2.11	2.6	1.27	3.52	2.49	0.99	0.78	1.18	2.21	2.09		
2.09	5 stall and putting to the table																										2.09	
Station-4 Flange Bending Process		OBSERVED TIMES																										
work element no	work element	t value	Standard deviation	Mean	k	n'	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20		
2.09	1 Taking panel from table	1.087	1.8315	0.05	615	0.9	0.36	0.48	0.51	0.98	1.29	2.33	0.89	3.55	0.99	1.58	1.91	2.11	3.21	3.36	1.59	1.69	2.28	4.01	2.61	2.61	2.09	
2.09	2 Putting panel into the press	0.993	2.374	0.05	306	1.57	1.12	3.24	0.98	0.87	2.31	2.87	3.64	1.69	0.91	2.64	3.67	3.67	1.85	2.53	2.67	4.13	2.59	1.82	3.51	3.51	2.09	
2.09	3 Standing on press button	0.148	1.5599	0.05	167	0.42	0.87	1.11	1.16	1.09	1.65	1.938	1.45	1.59	2.1	1.66	0.99	1.84	1.92	2.22	2.09	1.88	1.79	1.92	1.45	2.09		
2.09	4 Working the press	0.002	0.4929	0.05	0.02	0.492	0.492	0.492	0.495	0.496	0.493	0.492	0.496	0.492	0.492	0.493	0.493	0.492	0.492	0.493	0.493	0.495	0.495	0.495	0.495	0.492	2.09	
2.09	Taking the panel from press	0.861	2.74	0.05	173	1.25	3.42	1.66	2.28	2.64	2.98	4.12	3.66	1.24	1.68	1.94	2.59	2.97	2.61	3.16	4.1	2.98	2.71	3.14	3.67	2.09		
2.09	5 stall and putting to the table																										2.09	
Station-5 Flanges Ironing Process		OBSERVED TIMES																										
work element no	work element	t value	Standard deviation	Mean	k	n'	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20		
2.09	1 Taking panel from trolley	0.890	2.06	0.05	326	0.96	1.26	1.24	1.3	0.98	0.94	0.87	2.22	2.51	3.65	3.1	1.45	1.96	3.46	3.11	2.57	2.25	2.27	2.61	2.49	2.09		
2.09	2 Putting panel into the press	0.588	3.9115	0.05	39.5	3.11	2.99	3.27	3.21	5.23	3.46	4.02	3.91	4.87	4.42	3.99	4.1	3.99	4	3.97	3.51	4.27	3.96	3.96	3.96	2.09		
2.09	3 Standing on press button	0.176	1.521	0.05	23.3	1.32	1.69	1.54	1.67	1.66	1.98	1.59	1.36	1.34	1.64	1.49	1.48	1.44	1.61	1.67	1.59	1.32	1.29	1.41	1.33	2.09		
2.09	4 Working the press	0.019	4.3435	0.05	0.01	4.34	4.35	4.34	4.34	4.34	4.34	4.34	4.34	4.34	4.34	4.34	4.34	4.34	4.34	4.34	4.34	4.34	4.34	4.34	4.34	4.34	2.09	
2.09	Taking the panel from press	0.800	3.0215	0.05	122	2.98	3.21	3.45	4.27	4.56	4.29	5.26	5.31	5.6	3.49	4.16	5.45	4.99	4.58	4.76	5.26	3.91	5.16	4.66	3.89	5.61	2.09	
2.09	5 stall and putting to the table																										2.09	
Station-6 Border Bending Process		OBSERVED TIMES																										
work element no	work element	t value	Standard deviation	Mean	k	n'	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20		
2.09	1 Taking panel from table	0.652	1.627	0.05	280	0.72	1.12	1.12	0.98	0.94	0.71	1.26	1.98	1.87	2.65	2.12	0.81	2.71	1.46	2.19	2.61	1.75	1.56	1.63	2.11	2.09		
2.09	2 Putting panel into the press	0.554	2.5335	0.05	84.3	2.41	3.26	1.92	2.95	3.12	1.97	1.83	2.55	3.12	1.93	2.77	2.54	2.69	2.81	3.45	1.69	1.82	3.14	5.1	1.99	1.99	2.09	
2.09	3 Standing on press button	0.350	2.08	0.05	49.4	2.11	2.25	1.85	1.96	2.53	2.45	2.15	2.14	2.31	2.67	2.21	2.26	1.64	1.59	1.67	1.26	2.45	1.98	2.11	2.01	2.09		
2.09	4 Working the press	0.118	2.9185	0.05	237	2.94	2.98	3.02	2.96	2.64	2.65	2.67	2.99	3.02	3.03	2.94	2.98	2.96	2.94	2.94	2.94	2.95	2.93	2.95	2.94	2.09		
2.09	Taking the panel from press	0.580	4.8975	0.05	24.5	5.1	4.87	4.56	4.29	5.26	5.31	5.6	3.49	4.16	5.45	4.99	4.58	4.76	5.26	3.91	5.16	4.66	3.89	5.61	5.61	2.09		
2.09	5 stall and dragging to the table																										2.09	
Station-7 Hole Punching Process		OBSERVED TIMES																										
work element no	work element	t value	Standard deviation	Mean	k	n'	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20		
2.09	1 Taking panel from trolley																											

Table A.2. Time study form

TIME STUDY FORM				
Location: Machine Molding Factory				
Work Element No: 1				
Work Element Name: Bur Cutting Process				
Analyst Name: Esra Can				
Observation No	Observed time	Performance rating	Normal time	Allowances
1	1.50	88	1.32	16.17
2	1.80	88	1.58	16.17
3	1.32	88	1.16	16.17
4	1.26	88	1.11	16.17
5	1.98	88	1.74	16.17
6	1.50	88	1.32	16.17
7	2.04	88	1.80	16.17
8	1.56	88	1.37	16.17
9	1.62	88	1.43	16.17
10	2.34	88	2.06	16.17
11	1.32	88	1.16	16.17
12	2.40	88	2.11	16.17
13	1.20	88	1.06	16.17
14	1.38	88	1.21	16.17
15	2.34	88	2.06	16.17
16	1.92	88	1.69	16.17
17	1.50	88	1.32	16.17
18	1.98	88	1.74	16.17
19	1.86	88	1.64	16.17
20	1.74	88	1.53	16.17
Average Observation Time				1.73 seconds
Average Normal Time				1.52 seconds
Standard Time				1.76 seconds
Total Allowances Percentage				16.17%
Standard Deviation				0.372

**Research Article****An investigation on the flexural behaviour of RC beams wrapped with CFRP****İlknur Dalyan**^{a,*}  and **Bilge Doran**^b ^aDisaster and Emergency Management Presidency, Çankaya, Ankara, 06800, Turkey^bYıldız Technical University, Department of Civil Engineering, Esenler, Istanbul, 34220, Turkey

ARTICLE INFO

Article history:

Received 30 July 2020

Revised 24 September 2020

Accepted 08 October 2020

*Keywords:*Carbon fiber reinforced
polymer
Flexural behaviour
RC beams
Strengthening

ABSTRACT

In addition to conventional retrofitting of constructions, new technologies are rapidly being developed to withstand the external effects while sustaining an acceptable level of damage. Reinforced concrete structures strengthened with fiber reinforced polymer composite materials are becoming more and more widespread in structural applications due to their better mechanical properties, resistance to environmental influences, ease of application and light weight, as well as conventional methods of strengthening. In this study, strengthening technique as a methodology for externally bonded with carbon fiber reinforced polymer (CFRP) sheets to increase the flexural resistance of reinforced concrete beams has been investigated. For this purpose, eight reinforced concrete beams were produced considering different types of CFRP configuration and tested under four-point bending loading. The dimensions of the beams are 150×250×2600 mm and concrete cover of the first and second group of test beam are 20 mm and 40 mm, respectively. Finally, load-deflection behavior with the failure mechanism of the tested beams have been discussed and the effect of different schemes of strengthening on the flexural behavior has been evaluated.

© 2021, Advanced Researches and Engineering Journal (IAREJ) and the Author(s).

1. Introduction

The main purpose of strengthening methods is to bring the strength, ductility and rigidity of the element and/or system to desired level. In addition to conventional techniques, new techniques are being developed for the strengthening of structures and are widely used in the construction sector. In this field, one of these new techniques is the strengthening application with the use fiber reinforced polymer materials (FRP).

FRP materials have a wide usage area for strengthening of reinforced masonry and concrete construction elements depending upon high-grade properties such as light weight, ease of application in the area, high resistance of corrosion and high stiffness- and strength -to-weight ratio [1-6].

Bonding FRP materials is a new retrofitting or strengthening method for enhancing shear and flexural performance of existing reinforced concrete (RC) beams [7-9]. The effects of FRP materials which are externally bonded to the beams on flexure and/or shear capacity of

beams have been investigated by previous experimental and numerical studies [3-5], [9-24]. In these studies, it is seen that the FRP composites are utilized as externally epoxy-bonded reinforcement to enhance the structural strength and stiffness, ductility and seismic performance of the RC beams.

Dong et al. [4] performed an experimental investigation on both the shear-flexural and the flexural performance of strengthened RC beams by GFRP and CFRP sheets. For this purpose, fourteen identical beams were fabricated and tested to failure for evaluating how different strengthening configurations of GFRP and CFRP sheets affected RC beams. Results showed that the flexural strengthening plan was less adequate than the shear-flexural strengthening in increasing the ultimate strength. Besides, the numerical analysis was accomplished to evaluate the shear and bending capacities of the tested beams. Analysis results indicated that theoretical predictions were in good agreement with experimental ones. Siddiqui [10] investigated experimentally the effectiveness and

* Corresponding author. Tel.: +0-312-258-2323/2114; Fax: +0-312-258-21-02.

E-mail addresses: ilknur.dalyan@afad.gov.tr (I. Dalyan), doran@yildiz.edu.tr (B. Doran)

ORCID0000-0001-6436-7109 (I. Dalyan), 0000-0001-6703-7279 (B. Doran),

DOI: 10.35860/iarej.775557

This article is licensed under the CC BY-NC 4.0 International License (<https://creativecommons.org/licenses/by-nc/4.0/>).

adequacy of different FRP layouts in shear-flexural reinforcing of RC beams. Two groups of beams were fabricated, each group contained three beams as to be strong in shear and weak in flexure in the first group and to be weak in shear and strong in flexure in the second group. All beams were subjected to the same loading and the effectiveness of CFRP applications of different layouts was evaluated. It has been observed that CFRP sheets with U-shaped end anchored applied to the tensile surface of the beams are not only very effective in increasing the flexural capacity but providing sufficient deformation capacity as well. Nayak et al. [15] performed an experimental investigation of externally strengthened reinforced concrete beams by GFRP fabrics. For this purpose, reference RC beam and nine RC beams strengthened with GFRP fabrics in different ways were tested under two-point bending. They also developed a design proposal in addition to IS: 456-2000 to estimate the ultimate design resistance of RC beams which were strengthened with FRP fabric layers. For verification, the experimental flexural strength of strengthening RC beams were compared with the design flexural strength values calculated from the numerical model and the ACI 440-2R-08 design code. In addition, it was indicated that the flexural strength increased with an increase in number of fabrics for all strengthening schemes. Tautanji et al. [16] analyzed the failure mechanism of strengthened beams with FRP sheets. For this purpose, seven identical beams were fabricated and tested until failure. They also developed a moment displacement model and used it on beams which were tested in their study and other similar studies. The results showed that the ductility of strengthened beams by using externally bonded FRP sheets was significantly reduced compared to the reference beam. The outcomes also indicated that the proposed model validated the test results. Al-Amery and Al-Mahaidi [17] accomplished an experimental study to examine the flexural-shear strengthening of RC beams. The researchers tested an unstrengthened RC beam sample in addition to six strengthened RC beam samples which were strengthened in different scheme using CFRP straps and sheets. They indicated that coupling of CFRP sheets and straps improved the beam strength significantly. Also, it has been denoted that a more ductile behavior of RC beams can be obtained if the debonding failure is prevented. Esfahani et al. [18] studied the influence of reinforcement ratio on the flexural strength of reinforced concrete beams by utilizing CFRP layers. For investigation, twelve reinforced concrete beams were produced and four-point bending tests were performed on these beams. Test results and observations showed that an improvement of flexural stiffness and strength of strengthened beams were obtained. Reda et al. [19] performed an experimental study to investigate the behavior of RC beams which retrofitted with near surface

mounted (NSM) glass fiber reinforced polymer (GFRP) bars. For this purpose, a reference beam and ten strengthened RC beams were casted and tested under four-point bending. Beams were strengthened with straight GFRP bars and GFRP bars with bent end. In the experimental study, it was observed that the GFRP bars with bent ends increased the load carrying capacity and the concrete cover separation of the beams. Test results stated that load-deflection and GFRP strain-load values for strengthened beams with straight NSM bars were compatible the analytical prediction. Huang et al. [20] studied on the flexural behavior of RC beams strengthened with polyester FRP composite plates. For investigation, six strengthened RC beams and two reference beams were manufactured and tested under four-point bending. The steel reinforcement ratio and the thickness of PFRP plates were chosen by authors as experimental parameters. They also compared the experimental ultimate load values of the beams strengthened PFRP with the equations given in ACI 440.2R-08. Experimental results indicated that the ultimate load, the ductility and the deflection increased in PFRP retrofitted RC beams. Tashiri et al. [21] studied the failure modes, the load-deflection behavior and the crack propagation patterns of strengthened RC beams using RC jacketing and FRP strengthening. Twelve strengthened RC beams (six RC jacketed beams and six CFRP strengthened beams) and three unstrengthened beams were tested under three-point bending. The outcomes indicated that both strengthening techniques increase energy dissipation capacity and strength. Mahal et al. [22] explored the efficiency of different strengthening methods (plate strengthening and NSM bar-strengthening) of RC beams through experimental studies. They conducted four-point bending test with RC beams under fatigue and monotonic loading using unstrengthened and strengthened RC beams. The outcomes showed that the mid-span displacement for beams strengthened with NSM bars was higher than the one for beams strengthened with CFRP plates. Sharaky et al. [23] investigated the effectiveness of near-surface mounted (NSM) fiber reinforced polymer (FRP) reinforcement and axial stiffness on the strengthened beam failure modes and flexural capacities experimentally and numerically. An unstrengthened and seven different strengthened RC beams were tested under four-point loading. The experimental results indicated that the confinement significantly increased the ultimate load of the RC beams. Jawdhari et al. [24] performed an experimental study to examine the effectiveness and behavior of spliced CFRP rod panels (CRPs) using for strengthening of RC beams. Five strengthened RC beams which were strengthened in different scheme using CFRP rod panels and CFRP fabric besides an unreinforced RC beam were tested. The results showed that the failure load and ultimate load increased in strengthened RC beams.

The benefit of the CRP system was emphasized in the study.

In this study, the flexural behavior of CFRP wrapped RC beams have been studied experimentally. For this purpose, three different wrapping schemes have been considered. Based on the flexural behavior of eight RC beams tested under four-point bending loading, existing experimental data are also evaluated.

2. Experimental Program

An experimental research was conducted at Yıldız Technical University Structural Engineering Laboratory for investigating the influence of different CFRP schemes on flexural strengthening of RC beams. The test specimens consisted of eight RC beams and classified into two groups according to the concrete cover. The overall concrete cover thickness of beams groups 1 and 2 were 20 mm and 40 mm, respectively. Both in group 1 and 2, there were an unstrengthened RC beam (reference beam), and three RC beams that were differently bonded with CFRP sheets [6], [25].

2.1 Beam Specimens

A total of 8 RC beams were fabricated in two groups; each group containing 4 beams. The beams of the first group were fabricated to be concrete cover thickness as 20 mm, whereas beams of the second group were fabricated to be concrete cover thickness as 40 mm. All of the beams had identical cross-sectional dimensions, stirrups and longitudinal reinforcement. Beams had rectangular cross sections of 150x250 mm and span lengths of 2400 mm. All beams were reinforced with two 12 mm diameter steel bars on both sides of tension and compression. Shear reinforcement consisted of 8 mm diameter stirrups with 100 mm spacing. The geometry and steel reinforcement of the test beams are shown in Figure 1.

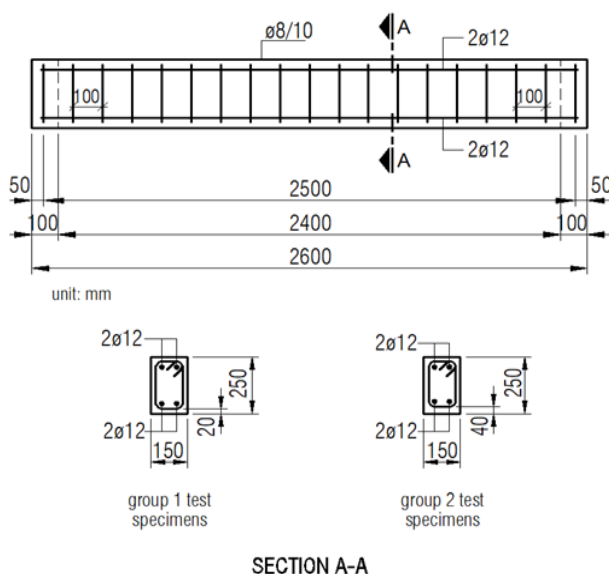


Figure 1. Geometry, steel reinforcement of beams tested in flexure

2.2 Material properties

2.2.1 Concrete and Steel

The concrete used in the fabrication of test beams was supplied from a ready-concrete company. During the concrete pouring, three cubes were taken to specify the 28-days compressive strength. Cube concrete compressive strength average of 25,1 MPa was converted to cylinder concrete compressive strength and determined as 20 MPa. Reinforcing steel bars with a yield strength of 420 MPa were used as stirrup and longitudinal reinforcement in beams. The properties of steel and concrete are summarized in Table 1.

2.2.2 Primer, Epoxy and CFRP sheet

Master Brace P3500, a two-component, low viscosity primer (Table 2), was used. Master Brace SAT4500 epoxy based adhesive (Table 3) which has a two-components, high strength, low viscosity and easy to apply was used to adhere the CFRP material to the beam surfaces. As for the strengthened material, Master Brace FIB 300/50 CFS unidirectional carbon fiber reinforced polymer (CFRP) was used. The properties of CFRP sheet is summarized in Table 4.

2.3 Strengthening procedure

In each group, one of the beams was not strengthened as a reference beam and the other three beams were strengthened using 0.17 mm thick unidirectional CFRP sheet in different arrangement. Firstly, the beam surfaces were sanded to smooth the rough beam surface and cleaned from the dust.

Primer material was applied to the beam surface to provide bonding with the beam surface and to obtain a smooth surface. Then epoxy adhesive was used and CFRP sheet was carefully bonded to the surface of the beam. The application of CFRP sheet bonding is given in Figure 2.

The strengthening schemes of the specimen RC beams are shown in Figure 3. These strengthening schemes are; Type 1: externally reinforcing with one layer of CFRP sheet on the bottom surface of the tension zone (150x2600 mm), Type 2: bonding of CFRP sheet to the bottom surface

Table 1. The properties of concrete and reinforcing steel properties

Material	Properties	Value
Concrete	Mean compressive strength (cube)	25.1 MPa
Concrete	Mean compressive strength (cylinder)	20.0 MPa
Reinforcing steel	Mean yield strength ($\phi 12$)	463.60 MPa
	Mean tensile strength ($\phi 12$)	572.07 MPa
	Mean yield strength ($\phi 10$)	518.03 MPa
	Mean tensile strength ($\phi 10$)	761.77 MPa
	Modulus of elasticity	2.10^5 MPa

Table 2. Technical properties of the primer

Compressive properties	Value	Tensile properties	Value
Yield strength	26.2 MPa	Yield strength	14.5 Mpa
Elastic modulus	670 MPa	Elastic modulus	717 Mpa
Ultimate strength	28.3 MPa	Ultimate strength	17.2 Mpa
Rupture strain	10 %	Rupture strain	40 %
Flexural properties	Value	Physical properties	Value
Yield strength	24.1 Mpa	Installed thickness	0.075 mm
Elastic modulus	595 Mpa	Density	1102 kg/m ³
Rupture strain	Large deformation with no rupture	Poisson ratio	0.48

Table 3. Technical properties of the epoxy adhesive

Compressive properties	Value	Tensile properties	Value
Yield strength	86.2 MPa	Yield strength	54 MPa
Elastic modulus	2620 MPa	Elastic modulus	3034 MPa
Ultimate strength	86.2 MPa	Ultimate strength	55.2 MPa
Rupture strain	5%	Rupture strain	3.5%
Flexural properties	Value		
Yield strength	138 MPa	Poisson ratio	0.40
Elastic modulus	3724 MPa		
Rupture strain	5%		

Table 4. Properties of CFRP sheet

Properties	Value
Modulus of Elasticity	230000 N/mm ²
Tensile strength	4900 N/mm ²
Thickness of design cross-section	0.166 mm
Weight	300 gr/m ²
Rupture strain	2.1%
Width	500 mm

and to both lateral side surfaces of 50 mm and 70 mm height to the tensile reinforcement including the concrete cover (250x2600 mm; 290x2600 mm), and Type 3: bonding of CFRP sheets to bottom surface and most of lateral sides (500x2600mm).

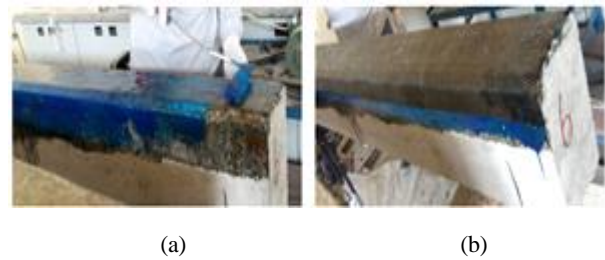
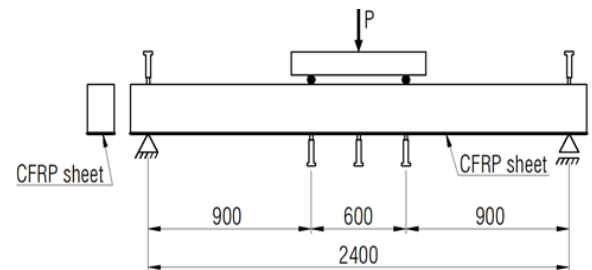
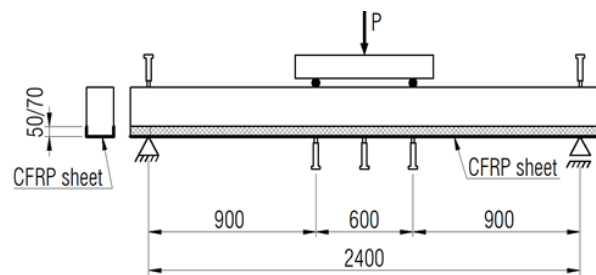


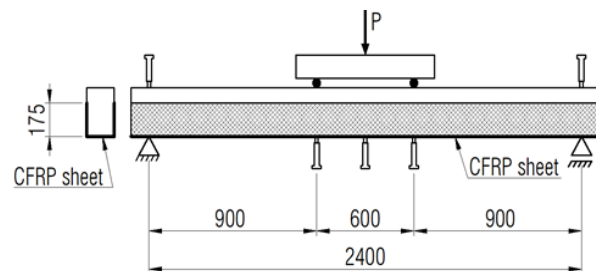
Figure 2. CFRP application of test beams a) epoxy adhesive application b) bonding of CFRP



(a) K20KG00 and K40KG00



(b) K20KG50 and K40KG70



(c) K20KGU and K40KGU

Figure 3. External strengthening arrangement of the beams tested in flexure

2.4 Experimental set up and instrumentation

As can be seen in Figure 4, eight rectangular concrete beam specimens were tested until failure under symmetric two point loads at spacing of $0.375 L = 900$ mm from the support points.

The load was then implemented using hydraulic actuator of 400 kN capacity with a constant loading rate of 2 mm/min. Vertical displacements at the midpoint of the beam, at the points where the load was implemented and on the supports were measured by linear variable differential transformers (LVDTs) [6].

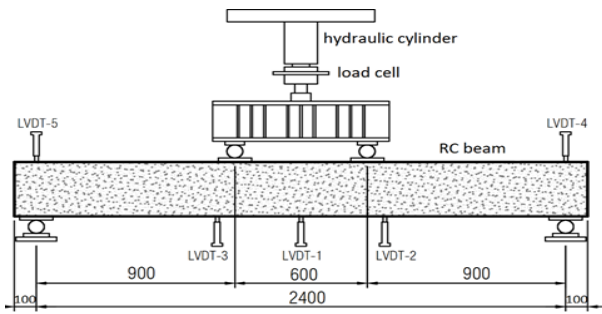


Figure 4. Experimental set-up

LVDTs were located with a measuring length of 300 mm in mid-span and with a measuring length of 100 mm at other points. The recording process was also continued until the completion of the test.

3. Results and Discussion

3.1 Failure modes and load-deflection behavior

In literature, failure mechanisms observed in RC beams externally strengthened with FRP sheets are flexural failure (concrete crushing and CFRP rupture), shear failure and FRP debonding [2], [26-30].

The most observed failure mechanisms are debonding of the CFRP material. Debonding usually takes place at the concrete-CFRP interface in a region of high stress concentration. Around the flexural and shear cracks and the ends of the CFRP material are debonding region [26]. Many researchers [2], [8], [26-30] refer to three different CFRP debonding failure modes: intermediate crack debonding, critical diagonal crack debonding and plate end debonding (Figure 5). Two different rupture modes are observed strengthened RC beams by CFRP composites: rupture prior to debonding and rupture after intermediate crack debonding [30].

3.1.1 Results for the 1st group of test specimens

Four RC beams (K20R, K20KG00, K20KG50, K20KGU) were loaded up to failure and their experimental ultimate loads are summarized in Table 5.

Failure modes observed for the first group of RC beams are indicated in Figure 6. The load-mid span deflection curves of the first group of RC beams are given in Figure 7. The ultimate load of the tested beams differed with the configuration of CFRP sheets.

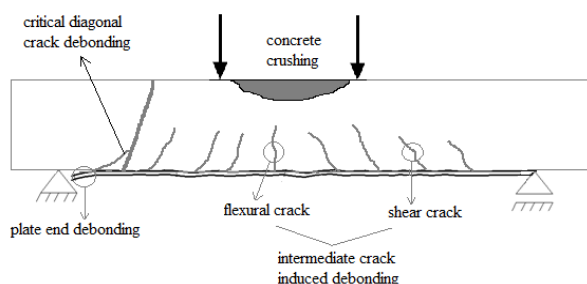


Figure 5. Failure modes of FRP [26], [27], [30]

Table 5. Test results of first group of beam

Beam ID	Ultimate Load (kN)	Failure Mechanism
K20R	51.97	Flexural failure
K20KG00	80.43	Debonding and rupture of CFRP
K20KG50	93.56	Debonding of CFRP
K20KGU	108.19	Debonding and rupture of CFRP

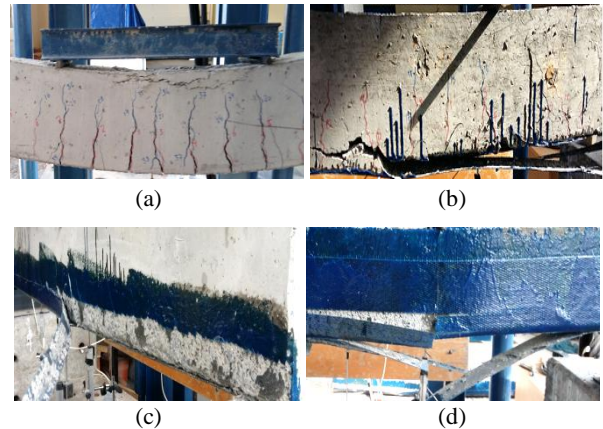


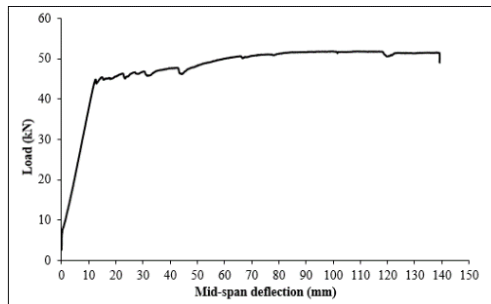
Figure 6. Failure modes of first group of RC beams
 a) flexural failure (K20R) b) intermediate crack induced debonding and rupture of CFRP (K20KG00) c) plate end debonding of CFRP (K20KG50) d) debonding and rupture of CFRP (K20KGU)

Typical flexural crack in the mid-span section and concrete crushing at compression region was observed for the reference beam (K20R) as seen in Figure 6(a). The first crack was monitored at the load level of 35 kN. During the experiment, mid-span deflection of 99.76 mm at the ultimate load level of 51.97 kN was monitored as seen in Figure 7(a). While the load remained constant, the displacement increased to a maximum value of 139.35 mm. In addition to the concrete crushing at compression region of beam, large crack widths and deflections were finally observed especially at the points where the load was applied. Finally, flexural failure was observed for the reference beam.

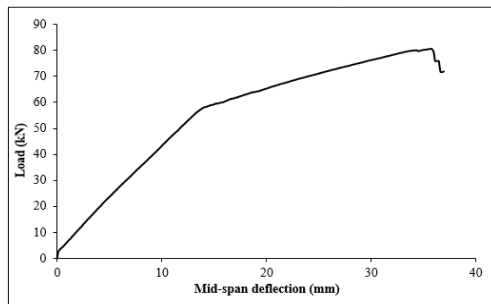
The failure mode for K20KG00 beam, intermediate crack induced debonding and rupture of CFRP was observed as seen in Figure 6(b). The first crack was monitored at the load level of 51 kN. During the experiment, mid-span deflection of 35.72 mm at the ultimate load level of 80.43 kN was monitored. Besides, maximum mid-span deflection of 36.93 mm was recorded as seen in Figure 7(b). The appearance of cracks significantly delayed and the displacement decreased compared to the reference beam (K20R). The mid-span deflection was decreased with the increase in the flexural rigidity due to the CFRP strengthening as expected.

Besides, the ultimate load was measured 54.8% greater than for the K20R beam.

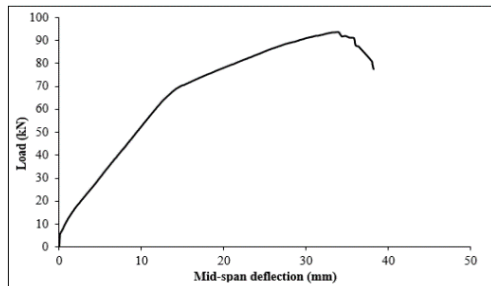
The failure mode for K20KG50, plate end debonding of CFRP was observed as shown in Figure 6(c). The first crack was monitored at the load level of 66 kN During the experiment, mid-span deflection of 33.88 mm at the ultimate load level of 93.56 kN was monitored. The maximum mid-span deflection of 39.15 mm was recorded as seen in Figure 7(c). The measured ultimate load of the test specimen K20KG50 is 80% greater than the K20R beam. Besides, maximum deflection at failure for K20KG00 and K20KG50 were lower than K20R beam.



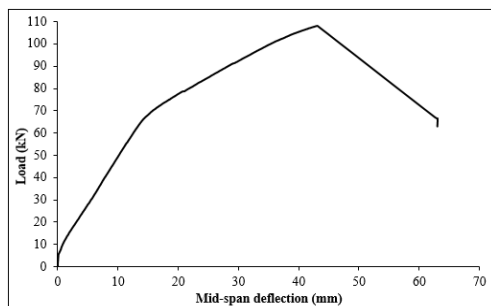
(a)



(b)



(c)



(d)

Figure 7. Load-mid-span deflection relationships for first group of beams a) K20R beam b) K20KG00 beam c) K20KG50 beam d) K20KGU beam

The failure mode for K20KGU, debonding and rupture of CFRP was observed as seen in Figure 6(d). During the experiment, mid-span deflection of 43.15 mm at the ultimate load level of 108.19 kN was monitored. Also, maximum mid-span deflection value of 63.07 mm was recorded as seen in Figure 7(d). The ultimate load was measured 108.2% greater than the reference beam (K20KG). The ultimate load increased by 34.5% and 15.6% compared to those of other strengthened beams (K20KG00 and K20KG50).

3.1.2 Results for 2nd group of test specimens

Experimental ultimate loads of four RC beams (K40R, K40KG00, K40KG70, K40KGU) are presented in Table 6.

Experimentally observed failure modes for the second group of RC beams are exhibited in Figure 8. The load-mid span deflection graphs of the second group RC beams are given in Figure 9.

Reference beam (K40R) was failed due to flexural crack at the mid-span section and crushing of concrete at the compression region as presented in Figure 8(a). The first crack was observed at the load level of 28 kN which was 0.8 times lower than K20R. During the experiment, mid-span deflection of 79.32 mm at the ultimate load level of 45.87 kN for K40R beam was monitored as seen in Figure 9 (a).

Table 6. Test results of the second group of beams

Beam ID	Ultimate Load (kN)	Failure Mechanism
K40R	45.87	Flexural failure
K40KG00	76.78	Debonding and rupture of CFRP
K40KG70	91.43	Debonding of CFRP
K40KGU	98.95	Debonding and rupture of CFRP



(a)

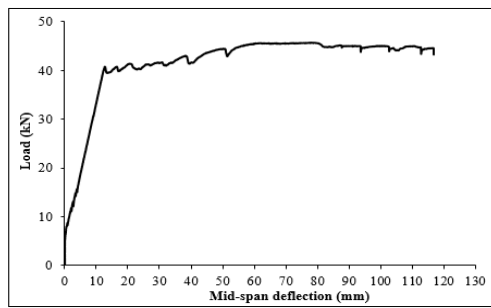
(b)



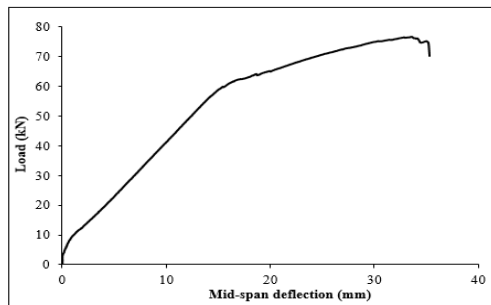
(c)

(d)

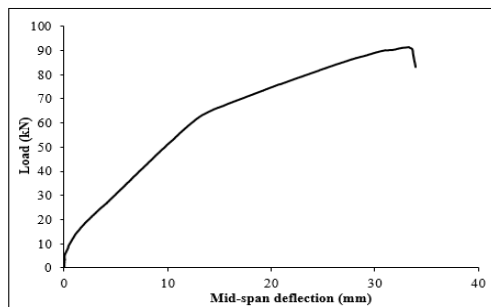
Figure 8. Failure modes of second group of RC beams a) crushing of concrete (K40R) b) intermediate crack induced debonding and rupture of CFRP (K40KG00) c) plate end debonding of CFRP (K40KG70) d) debonding and rupture of CFRP (K40KGU)



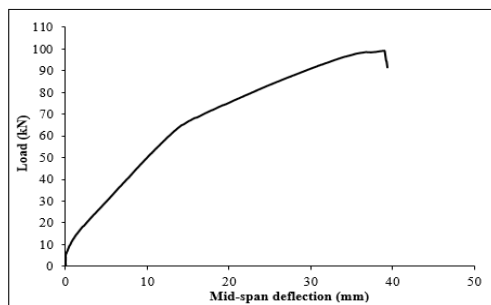
(a)



(b)



(c)



(d)

Figure 9. Load-mid-span deflection relationships for second group of beams a) K40R beam b) K40KG00 beam c) K40KG70 beam d) K40KGU beam

While the load remained constant, the displacement increased to a maximum of 117.13 mm. The large crack widths and deflections were observed until the failure. Crushing of concrete in the pressure zone of K20R and K40R beams was one of the reasons for the collapse these beams.

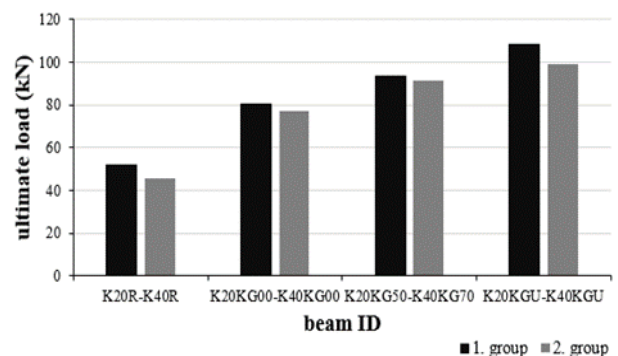
The strengthened beam K40KG00 failed due to intermediate crack induced debonding and rupture of CFRP at load of 76.78 kN as shown in Figure 8(b). The first crack was monitored at the load level of 46 kN during

the experiment, mid-span deflection of 34.63 mm at the ultimate load level of 76.78 kN was monitored. Besides, maximum mid-span deflection of 36.51 mm was recorded as seen in Figure 9(b). The ultimate and cracking loads of K40KG00 were much lower than K20KG00 as expected. The ultimate load of K40KG00 increased by 67.4% compared to that of the reference beam (K40R).

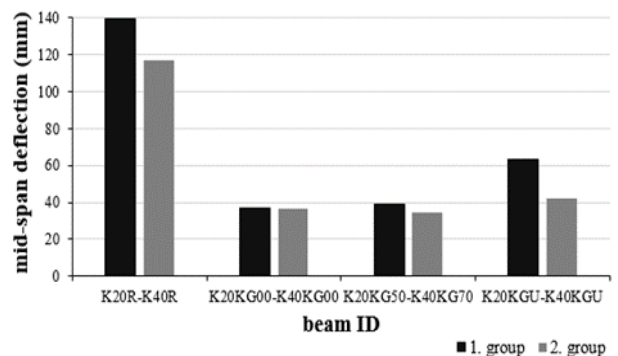
The failure mode for K40KG70, plate end debonding of CFRP was observed as shown in Figure 8(c). During the experiment, mid-span deflection of 33.82 mm at the ultimate load level of 91.43 kN was monitored. Maximum mid-span deflection of 34.56 mm was recorded as seen in Figure 9(c). First crack was monitored at the load level of 61 kN. The ultimate load of the test specimen K40KG70 increased by 99.3% compared to K40R.

K40KGU failed by debonding and rupture of CFRP at beam surface (Figure 8(d)). The ultimate load of the beam K40KGU was measured as 98.95 kN which was 115.7% greater than the reference beam K40R. The mid span deflection value was observed 40.04 mm at this load level. Also, maximum mid-span deflection value of 41.82 mm was recorded as seen in Figure 9(d). The ultimate load for K40KGU increased by 28.9% compared to that of K40KG00. Maximum deflection at failure for K40KGU was higher than the other strengthened beams (K40KG00 and K40KG70).

The effect of strengthening on the load and mid-span deflection for the experiment beams are given in Figure 10.



(a) Ultimate load



(b) Mid-span deflection

Figure 10. The effect of strengthening on the load and mid-span deflection

The ultimate load and maximum mid-span deflection at failure for K20R beam were more than the K40R beam. Although the ultimate load for K20KG00 beam was more than K40KG00, the maximum mid-span deflection at failure was nearly the same.

The measured values of ultimate load and maximum mid-span deflection were bigger in K20KG50 compared to K40KG70 and also bigger in K20KGU compared to K40KGU beams.

The first cracks in the K20KGU and K40KGU beams were observed at higher load levels that were compared to the other strengthened beams. Although the cost of strengthening was higher for the K20KGU and K40KGU beams than the K20KG50 and K40KG70 beams, there was no difference in load carrying performance significantly.

The ultimate loads for test beams having 20 mm concrete cover (K20R, K20KG00, K20KG50, K20KGU) increased when compared to the test beams having 40 mm concrete cover (K40R, K40KG00, K40KG70, K40KGU); an increase of specimens' ultimate load by 13.3%, 4.8%, 2.3% and 9.3% respectively. The difference in ultimate load (13%) between the reference beams K20R and K40R was caused by the effective depth values that are chosen as 216 mm and 196 mm respectively. This difference was considerably reduced by external bonding of CFRP sheet to bottom surface and to both lateral side surfaces of 50 mm and 70 mm height.

The mid-span deflection in strengthened beams decreased when compared to reference beams by 73.5%, 71.9%, 54.7% for K20KG00, K20KG50, K20KGU and 68.8%, 70.5%, 64.3% for K40KG00, K40KG70 K40KGU, respectively.

Considering the ultimate loads for test beams, this experimental study indicated that the strengthening scheme Type 2 is almost as effective as the Type 3.

4. Conclusions

In the context of this study, the flexural behavior of CFRP wrapped RC beams under four-point bending loading were investigated experimentally. The below mentioned conclusions can be achieved from this study:

- Experimental tests confirmed that the unstrengthened beams failed flexural failure and strengthened beams with CFRP sheets failed in two different modes; CFRP debonding and CFRP rupture.
- Ultimate loads for experimental beams wrapped with CFRP sheets are substantially increased compared to the reference beams.
- Mid-span deflection in strengthened beams is decreased in comparison with the reference beams. Externally bonded CFRP sheet leading to a decrease in flexural deformation as expected.

- It is concluded that the strengthening schemes Type 2 and Type 3 compared to Type 1 are an effective and successful method for obtaining a better flexural performance. It is observed that the strengthening scheme, which includes the concrete cover of RC beams, increases the flexural performance. Due to the amount of material used in strengthening scheme Type 3, this increase, when compared with Type 2 provides less performance-cost value. The strengthening scheme Type 2 is more effective than Type 1 in terms of flexural capacity and it is more economical than Type 3.

Acknowledgment

This work supported by Yildiz Technical University under Scientific Research Projects (project no: 2016-05-01-DOP06), Turkey.

Declaration

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article. The author(s) also declared that this article is original, was prepared in accordance with international publication and research ethics, and ethical committee permission or any special permission is not required.

Author Contributions

Authors planned the experiment. İ. Dalyan conducted experiments of beams. B. Doran evaluated experimental methods, measuring techniques and results. İ. Dalyan wrote the article. Editing and proofreading have been done by B. Doran.

References

1. Tanarşlan, H.M., *Repairing and strengthening of earthquake damaged RC beams with CFRP strips*. Magazine of Concrete Research, 2010. **62**(5): p. 365-378.
2. Niu, H. and Z. Wu, *Numerical Analysis of Debonding Mechanisms in FRP-Strengthened RC Beams*. Computer-Aided Civil and Infrastructure Engineering, 2005. **20**(5): p. 354-368.
3. Jankowiak, I., *Analysis of RC Beams Strengthened by CFRP Strips-Experimental and FEA study*. Archives of Civil and Mechanical Engineering, 2012. **12**: p. 376-388.
4. Dong, J., Q. Wang, and Z. Guan, *Structural Behaviour of RC Beams with External Flexural and Flexural-shear Strengthening by FRP Sheets*. Composites: Part B, 2013. **44**: p. 604-612.
5. Al-Rousan, R.Z., *Effect of CFRP Schemes on the Flexural Behavior of RC beams Modeled by Using a Nonlinear Finite-element Analysis*. Mechanics of Composite Materials, 2015. **51**(4): p. 437-446.
6. Ayaydın, M., İ. Dalyan and B. Doran. *Nonlinear Analysis of RC Beams Wrapped with CFRP*. in ICOCEE2018: Izmir. p. 1-12.

7. ACI 440.2R-08, *Guide for the Design and Construction of Externally Bonded FRP Systems for Strengthening Concrete Structures*. American Concrete Institute, 2008, U.S.A.
8. Kotynia, R., *Bond Between FRP and Concrete in Reinforced Concrete Beams Strengthened with Near Surface Mounted and Externally Bonded Reinforcement*. Construction and Building Materials, 2012. **32**: p. 41–54.
9. Akram, A., R. Hameed, Z.A. Siddiqi, M.R. Riaz, and M. Ilyas, *Finite Element Modeling of RC Beams Strengthened in Flexure using FRP Material*. Arabian Journal for Science and Engineering, 2014. **39**: p. 8573–8584.
10. Siddiqui, N.A., *Experimental Investigation of RC Beams Strengthened with Externally Bonded FRP Composites*. Latin American Journal of Solids and Structures, 2009. **6**: p. 343-362.
11. Barris, C., L. Torres, A. Turon, M. Baena, and A. Catalan, *An Experimental Study of the Flexural Behaviour of GFRP RC Beams and Comparison with Prediction Models*. Composite Structures, 2009. **91**: p. 286-295.
12. Godat, A., Z. Qu, X.Z. Lu, P. Labossière, L.P. Ye, and K.W. Neale, *Size Effects for Reinforced Concrete Beams Strengthened in Shear with CFRP Strips*. Journal of Composites For Construction, 2010. **14**(3): p. 260-271.
13. Huang, J., *Finite Element Modeling (FEM) of GFRP Bar Reinforced Concrete Beam: Flexural Behavior*. Advanced Materials Research, 2011. Vols. 255-260: p. 3114-3118.
14. Aktan, S., K. Polat, H.O. Köksal, B. Doran, and B. Deliktaş, *Numerical Modeling for the Flexural Behaviour of Reinforced Concrete Beams Strengthened with FRP Sheets*. in ACE 2014: İstanbul. p. 271-271.
15. Nayak, A.N., A. Kumari, and R.B. Swain, *Strengthening of RC Beams Using Externally Bonded Fibre Reinforced Polymer Composites*. Structures, 2018. **14**: p. 137-152.
16. Toutanji, H., L. Zhao, and Y. Zhang, *Flexural Behavior of Reinforced Concrete Beams Externally Strengthened with CFRP Sheets Bonded with an Inorganic Matrix*. Engineering Structures, 2006. **28**: p. 557-566.
17. Al-Amery, R. and R. Al-Mahaidi, *Coupled Flexural-Shear Retrofitting of RC Beams Using CFRP Straps*. Composite Structures, 2006. **75**: p. 457-464.
18. Esfahani, M.R., M.R. Kianoush, and A.R. Tajari, *Flexural Behaviour of Reinforced Concrete Beams Strengthened by CFRP Sheets*. Engineering Structures, 2007. **29**: p. 2428-2444.
19. Reda, R.M., I.A. Sharaky, M. Ghanem, M.H. Seleem, and H.E.M. Sallam, *Flexural Behaviour of RC Beams Strengthened by NSM GFRP Bars Having Different End Conditions*. Composite Structures, 2016. **147**: p. 131-142.
20. Huang, L., L. Zhao and L. Yan, *Flexural Performance of RC Beams Strengthened with Polyester FRP Composites*. International Journal of Civil Engineering, 2018. **16**: p. 715-724.
21. Tahsiri, H., O. Semidi, A. Khaloo and E.M. Raisi, *Experimental study of RC jacketed and CFRP strengthened RC beams*. Construction and Building Materials, 2015. **95**: p. 476-485.
22. Mahal, M., B. Täljsten, and T. Blanksvärd, *Experimental performance of RC beams strengthened with FRP materials under monotonic and fatigue loads*. Construction and Building Materials, 2016. **122**: p. 126-139.
23. Sharaky, I.A., M. Baena, C. Barris, H.E.M. Sallam, and L. Torres, *Effect of axial stiffness of NSM FRP reinforcement and concrete cover confinement on flexural behavior of strengthened RC beams: Experimental and numerical study*. Engineering Structures, 2018. **173**: p. 987-1001.
24. Jawdhari, A., A. Peiris, and I. Harik, *Experimental study on RC beams strengthened with CFRP rod panels*. Engineering Structures, 2018. **39**(5): p. 826-841.
25. Dalyan, İ., *Flexural Performance Evaluation of RC Beams Differently Strengthened with Fibre Reinforced Polymer*. in Civil Engineering 2019, Yildiz Technical University: Turkey. p. 185. (in Turkish)
26. Aram, M.R., C. Czaderski, and M. Motavalli, *Debonding Failure Modes of Flexural FRP-Strengthened RC Beams*. Composites Part B: Engineering, 2008. **39**(5): p. 826-841.
27. Oehlers, D.J., S.M. Park, and M.S. Mohamed Ali, *A structural engineering approach to adhesive bonding longitudinal plates to RC beams and slabs*. Composites: Part A: applied science and manufacturing, 2003. **34**(2003): p. 887-897.
28. Oehlers, D.J., P. Visintin, and W. Lucas, *Flexural Strength and Ductility of FRP-Plated RC Beams: Fundamental Mechanics Incorporating Local and Global IC Debonding*. Journal of Composites for Construction, 2016. **20**(2): p. 1-13.
29. Bhatti, A.Q. and N. Kishi, *Control of FRP Debonding in Strengthened RC Beams*. Arabian Journal for Science and Engineering, 2012. **37**(8): p. 2103–2112.
30. Wang, W-W., J-G. Dai and K.A. Harries, *Intermediate crack-induced debonding in RC beams externally strengthened with prestressed FRP laminates*. Journal of Reinforced Plastics and Composites, 2013. **32**(23): p. 1842–1857.



Research Article

Comparison of the effect of foundation analysis methods on structural analysis results of tall buildings

Ömer Özer ^{a,*}  and S. Bahadır Yüksel ^a 

^aKonya Technical University, Faculty of Engineering and Natural Sciences, Department of Civil Engineering, 42130, Konya, Turkey

ARTICLE INFO

Article history:

Received 23 September 2020

Revised 17 December 2020

Accepted 02 January 2021

Keywords:

Coefficient of subgrade reaction

Fixed base method

Pseudo-coupled method

Soil structure interaction

Winkler method

ABSTRACT

In structural analysis, there are serious interactions between soil-foundation and structure. The fixed base analysis method ignores this interaction but many analysis methods have been developed that take into account the soil-structure interaction (SSI). This paper revealed the effects of SSI analysis methods on tall buildings analysis results on soft soil sites. 6 different structural models with 20 story buildings and two different coefficients of subgrade reactions were analyzed with three different analysis methods which are fixed base method, Winkler method and pseudo-coupled method. As a result of all the analysis, 6 different structural models compared and discussed in terms of structural period, displacements, lateral loads, column, shear wall and foundation reactions and structural economy. It has been observed that the structure may remain on the unsafe side in the analysis made with the fixed base analysis method.

© 2021, Advanced Researches and Engineering Journal (IAREJ) and the Author(s).

1. Introduction

The production of the structures starts with the projecting process and is completed with the application process. In addition to being compatible with the architectural project at the project planning stage, the most important principle that civil engineers should consider is modeling as close as possible to the reality. When modeling is done, it is necessary to take into consideration both the superstructure conditions of the building and the geotechnical conditions of the region and the ground. In this context, it is very important that the ground survey reports should be evaluated correctly and used completely in the project.

In the analysis of the buildings, the soil structure interaction (SSI) is widely ignored. However, when we look at the studies in the literature, the local soil properties have significant effects on the behavior of the structure, under both vertical and horizontal loads [1]. Under vertical static loads, there is a tendency of displacement and rotation effects on the foundation of the structure and a tendency to move differently between the structure and the foundation under horizontal dynamic

loads. According to this situation, fixed base foundation acceptance is away from realism and analysis methods which take into account SSI is necessary.

Winkler method [2] is widely used in SSI models. However, in Winkler method, a realistic calculation of foundation shape changes cannot be made [3]. In addition to the Winkler method, the horizontal springs have also been introduced to achieve more realistic results, but additional difficulties have arisen in this method [4].

The normal stress between the foundation and the ground varies depending on the stiffness of the foundation system and the load distribution of the structure and the soil group [5]. The equivalent coefficient of subgrade reaction method, idealizing the difference of deformations arising from these differences, has been proposed by Vallabhan and Daloglu [6].

2.2.1.5 of Turkish Earthquake Code 2007 which was previous earthquake code required SSI analysis for soil group C and D [7]. General belief is that the period of the building will increase in the SSI solution, and the earthquake load affected by the building will also decrease, and the building performance will be improved.

* Corresponding author. Tel.: +90-536-516-1941 ; Fax: +90-332-238-2727.

E-mail addresses: omer@anahartarmuhendislik.com (Ö. Özer), sbyuksel@ktun.edu.tr (S.B. Yüksel)

ORCID: 0000-0002-5126-6832 (Ö. Özer), 0000-0002-4175-1156 (S.B. Yüksel),

DOI: 10.35860/iarej.799055

This article is licensed under the CC BY-NC 4.0 International License (<https://creativecommons.org/licenses/by-nc/4.0/>).

For the same reason, Turkish Building Earthquake Code 2018 16C.1.2 suggests that SSI can be neglected in to stay on the safe side. On the other hand, studies in the literature have shown that SSI do not only affect structural period but also other parameters. Fatahi et al. [8] investigated that the structural displacements and inter-storey drifts caused by SSI are larger than the corresponding values while only local site effect is included. The numerical results clearly indicate that the structural displacements and inter-storey drifts caused by SSI are larger than the corresponding values while only local site effect is included. Tomeo et al. [9] emphasized that SSI effects are important for soft soils. Moghaddasi et al. [10] emphasized that major earthquakes has highlighted the possibility of detrimental effects or increase in the structural response due to SSI. Çelebi et al. [11] stated that dynamic response is more pronounced for resonance case, when the frequency content of the seismic ground motion is close to that of the SSI system.

The purpose of this study is to show that the SSI affects many parameters other than the structural period only, and to show that the fixed base acceptance is not always on the safe side.

2. Material and Method

In this paper, 6 different structural models with 20 story (Figure 1-2) and two different soil groups (Table 1) were analyzed with three different analysis methods which are fixed base method, the Winkler method and the pseudo-coupled method.

Within the soil group D and C, the coefficients of subgrade reaction were selected as 12753 kN/m³ and 17167 kN/m³, respectively. The other important parameter of the soil group is “allowable bearing value” taken as 98 kN/m², and 147 kN/m².

In order to examine the effect of soil groups on foundation analysis methods, two different soil groups were examined. In this way, it is aimed not only to compare the

fixed base and SSI, but also to show how the solutions can differ according to soil groups.

In our country, the PGA value can rise up to 0.780 in the eastern Anatolian region. However, in our study, we wanted to choose a region where the population is high in addition to high earthquake acceleration (Table 2). For this reason, we preferred the ground acceleration of the Gölcük region, which caused great damage in 1999 [12].

Table 1. The properties of soil groups according to TBEC (2018) [13]

Soil Group	Description of Soil Group
C	Highly tight sand gravel and hard clay layers or with cracked weak rock
D	Medium-firm - firm sand, gravel or solid clay

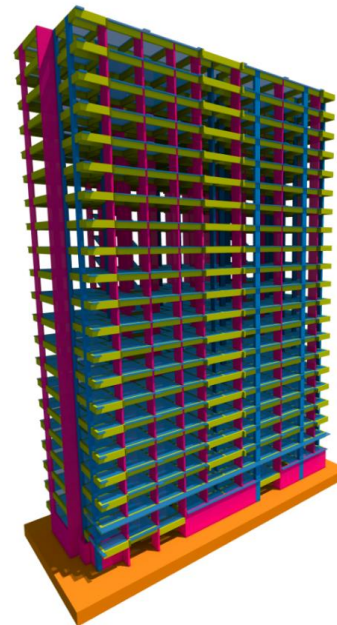


Figure 1. The perspective view of 20-storey building

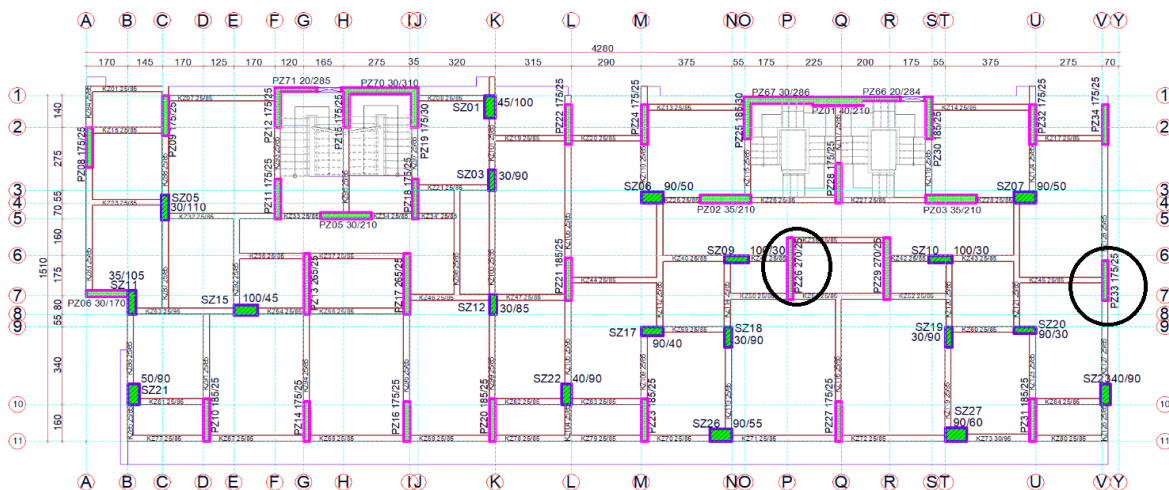


Figure 2. The floor plan of all structural models

Table 2. Population affected by Gölcük earthquake [12]

City	Center	Villages	Total
Bilecik	116004	76056	192060
Bolu	265052	287970	553022
Bursa	1484838	473691	1958529
Eskişehir	518643	142200	660843
İstanbul	8506026	692783	9198809
Kocaeli	629333	548046	1177379
Sakarya	331431	400369	731800
Tekirdağ	358878	208518	567396
Yalova	110106	53810	163916
Zonguldak	239186	373536	612722
Total	12.559.497	3.256.979	15.816.476

Table 3. Seismic parameters of analysis model

Structure Location	Kocaeli Gölcük
Latitude	40.720382
Longitude	29.811135
S _s	1.683
S ₁	0.399
S _{D1}	0.599
S _{DS}	2.02
PGA	0.687
PGV	43.509

Analysis model’s seismic location parameters are given in Table 3 according to TBEC 2018 [13]. The meanings of the symbols in Table 3 are given below:

- S_s: Short period map spectral acceleration coefficient
 - S₁: 1 second period map spectral acceleration coefficient
 - S_{d1}: 1 second period design spectral acceleration coefficient
 - S_{ds}: Short period design spectral acceleration coefficient
 - PGA: Peak ground acceleration (g)
 - PGV: Peak ground velocity (cm/sec)
- Analysis model’s general seismic design parameters are

given in Table 4.

Analysis methods can be summarized as follows:

Fixed base method:

In this method, it is assumed that the bottom floor columns and the shear walls were connected to the base in an infinitely rigid manner. Therefore, it was assumed that the foundation was not affected by the structure, and the structure was not affected by the foundation. In this case, foundation and structure are analyzed independently of each other.

Winkler method:

It was assumed that the shape of the springs changed only when loaded directly and formed a counter reaction, but each spring was considered to be independent of neighboring springs. That is, the ground was modeled by independent linear springs defined by the coefficient of subgrade reaction, but the continuity of the (soil) space was not taken into consideration. As a result, the ground was considered as a completely discontinuous environment. This inadequacy in the Winkler method has attracted the attention of researchers, and for this reason, the criticism and the recommendation for the improvement of the classical Winkler method were widely published in the literature [14].

Pseudo-coupled method:

In this method, the foundation was divided into three parts and three different coefficients of subgrade reaction selected for each part. In this method, the lower coefficient of subgrade is used in the structure core, while the coefficient of subgrade is increased as it moves to the foundation corners. The basis of the pseudo-coupled practice is shown in Figure 3 [15]. In this study, this method will be used as the third foundation analysis method.

Table 4. General design parameters of analysis model

Analysis Model Parameters			
Number of Story	21	Beam Dimensions	30 cm x 85 cm
Story Height	3 m	Beam Reinforcement - Confinement	6Φ22 - Up
Structure Dimensions (X-Y)	47 m x 20.8 m		6Φ22 - Down
Structure Height	63 m	Slab Height	12-15 cm
Structure Height Classes	BYS 2	Slab Dead / Live Load	2.06 / 3.43 kN/m ²
		Building Risk Category	BRC 3
Response Modification Coefficient R	8		
Overstrenght Factor D	2.5		
Concrete	C40	Analysis Method	Strength Based Design
Reinforcement Material	B420C	Ductility	High
Seismic Design Category	DTS 1	Total Weight of the Structure	168516 kN

In the study, sizes and reinforcements for columns and shear walls are shown in Table 4 and Table 5 (columns and shear walls of the same dimensions have the same number of reinforcement).

Each analysis method used in this study has serious effects on the analysis result of the structure. Figure 4 shows the effect of the foundation analysis methods on the foundation deformation under the same loading.

Table 5. Column dimensions and reinforcements bars

Column	b	h	Major	Minor	Confinement
SZ01	45	100	8Ø20	10Ø20	ø10/10/10
SZ03	30	90	4Ø20	6Ø20	ø10/12/8/10
SZ05	30	110	4Ø20	8Ø20	ø10/10/10
SZ07	90	50	8Ø20	10Ø20	ø10/20/10/10
SZ09	100	30	4Ø20	6Ø20	ø10/12/8/10
SZ11	35	105	6Ø20	6Ø20	ø10/10/10
SZ12	30	85	4Ø20	6Ø20	ø10/15/10/10
SZ17	90	40	6Ø20	6Ø20	ø10/20/10/10
SZ26	90	55	10Ø20	8Ø20	ø10/20/10/10
SZ27	90	60	10Ø20	8Ø20	ø10/20/10/10

Table 6. Shear wall dimensions and reinforcements bars

Shear Wall	b	h	Vertical Bar			Confinement		
PZ02	35	210	20	Ø	14	ø8	/	11
PZ05	30	210	36	Ø	14	ø8	/	13
PZ06	30	170	32	Ø	18	ø8	/	13
PZ08	25	175	18	Ø	14	ø8	/	16
PZ10	25	185	28	Ø	14	ø8	/	16
PZ13	25	265	30	Ø	14	ø8	/	16
PZ19	30	175	24	Ø	14	ø8	/	13
PZ25	30	185	22	Ø	14	ø8	/	13
PZ26	25	270	32	Ø	14	ø8	/	16
PZ66	20	284	42	Ø	14	ø12	/	10
PZ70	30	310	46	Ø	16	ø12	/	14
PZ71	25	285	46	Ø	14	ø12	/	12
PZ01	40	210	20	Ø	14	ø12	/	13
PZ67	30	286	56	Ø	16	ø12	/	12

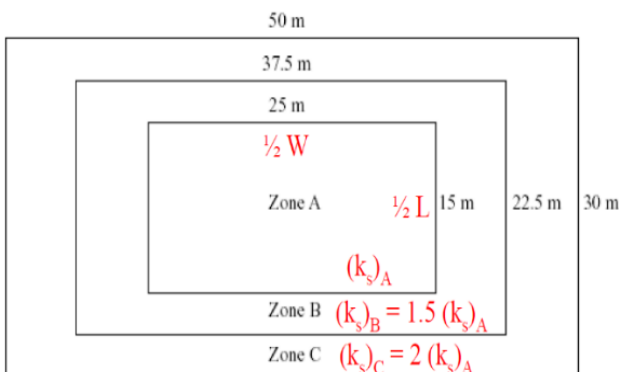


Figure 3. The coefficient of subgrade reaction in the regularly formed foundation area in the pseudo-coupled method [15]

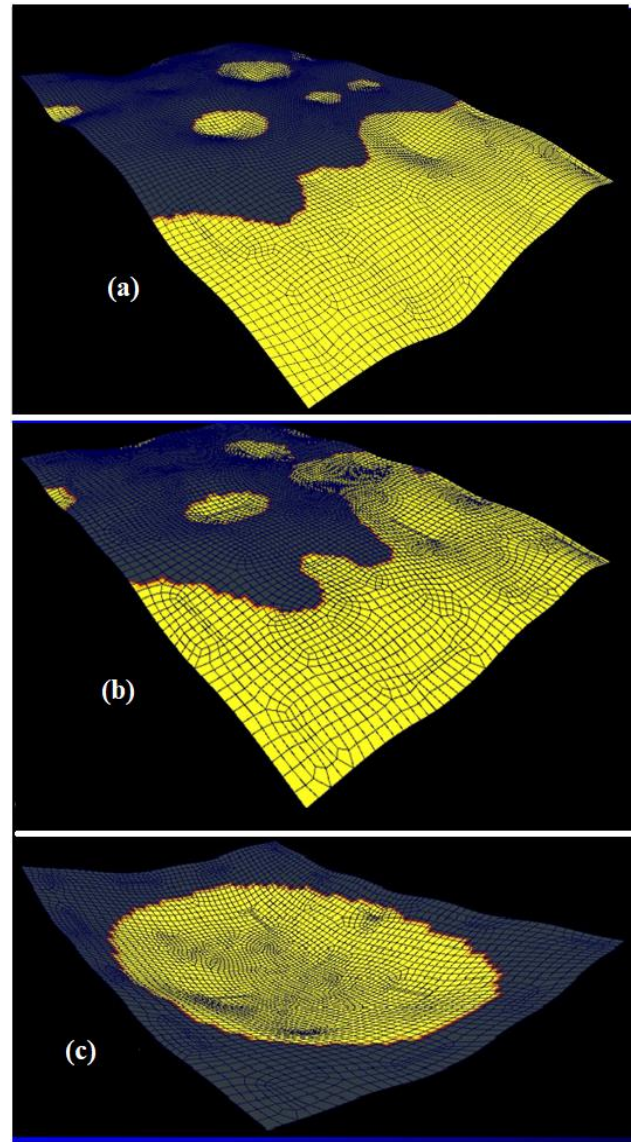


Figure 4. Effect of the foundation analysis methods on the foundation deformation under the same loading
a) Fixed base, b) Winkler, c) Pseudo-coupled

3. Research Results

As in many studies [16-18], it has been observed that the period of the structure increases in the acceptance of the SSI. Similarly, in this study, it has been observed that the building period increases in the SSI solutions (Figure 5). Similar to previous studies [9], when the top floor displacement values were examined, it was seen that the total displacement of the structure increased in SSI solutions (Figure 6), when the top floor earthquake loads and torsional moments are examined, it has been seen that the fixed base acceptance gets higher values (Figure 7 - 8).

In this study, the analysis results of all columns were examined in detail under headings of column capacity ratios, axial loads, earthquake loads and bending moments. The results of the research showed that the columns in the fixed base analysis system were subjected to higher axial loads, on the other hand the columns in the SSI systems were subjected to higher moments (Figure 9).

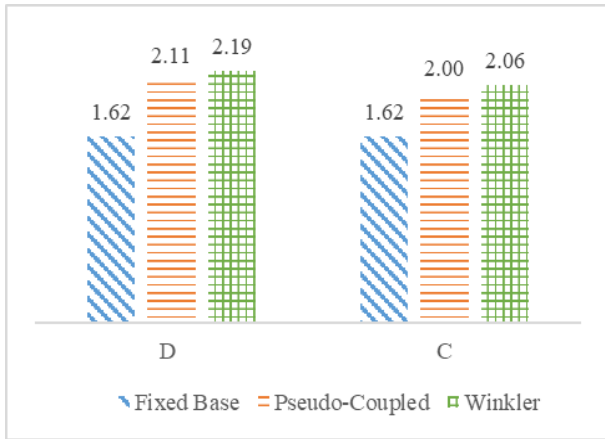


Figure 5. The structural periods (sec)

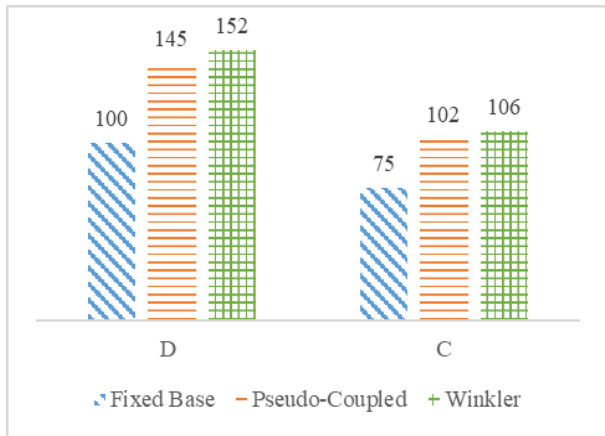


Figure 6. Relative top floor displacement (mm)

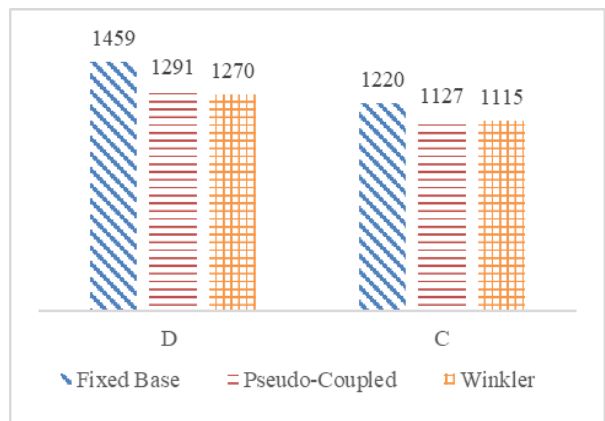


Figure 7. Top floor horizontal force (kN)

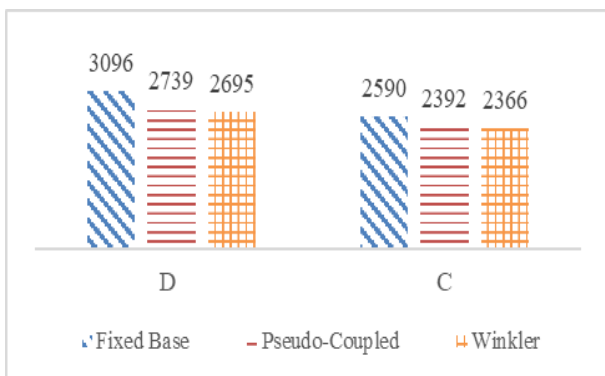


Figure 8. Top floor torsional moment (kNm)

N_{dMax} : Maximum axial force acting on column
 N_{max} : Maximum column axial force capacity
 M_d : Column bending moment
 M_r : Column bending moment capacity.

In this study, analysis results of PZ26 and PZ33 shear walls which was shown in Figure 2 was evaluated. Maximum axial loads of PZ26 and PZ33 shear walls are shown in Figure 10.

However, when the earthquake loads and M_3 bending moments were examined, it was seen that the fixed base acceptance received higher results than SSI acceptance (Figure 11-12).

In this study, raft foundation was preferred. The dimensions of the raft foundation were 4700 x 2080 cm and the foundation height were 170 cm for 20-story buildings.

Another subject that was compared in this study is soil stresses which were compared in detail for three different analysis methods and two different soil groups. When the research results were examined, it was seen the analysis results are varied for all analysis methods. The results of the analysis are shown graphically as minimum, average and maximum stress (Figure 13).

When the bending moments are examined, the biggest bending moment in the X direction was seen in the fixed base model, but when the Y direction is examined the biggest bending moment was seen in the pseudo-coupled model (Figure 14).

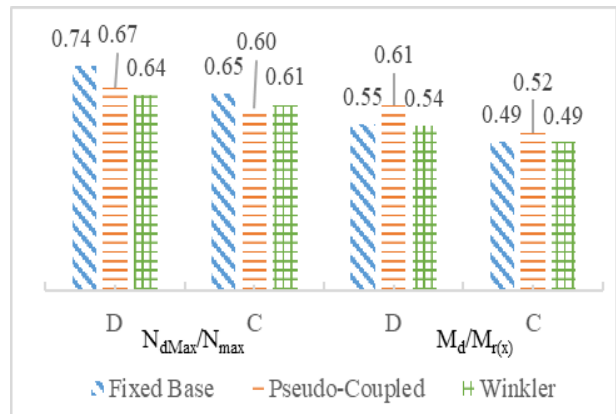


Figure 9. Ground floor columns average capacity ratios

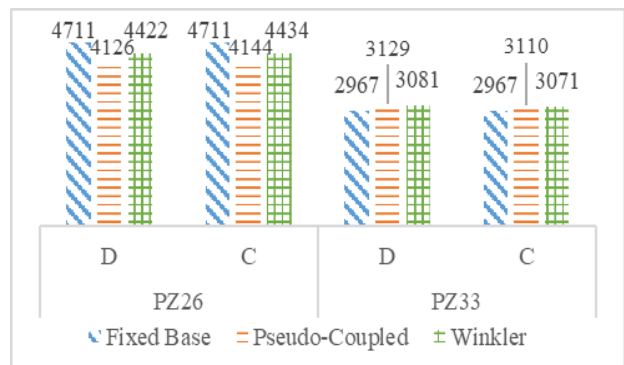


Figure 10. Axial load values of PZ26 and PZ33 shear walls (kN)

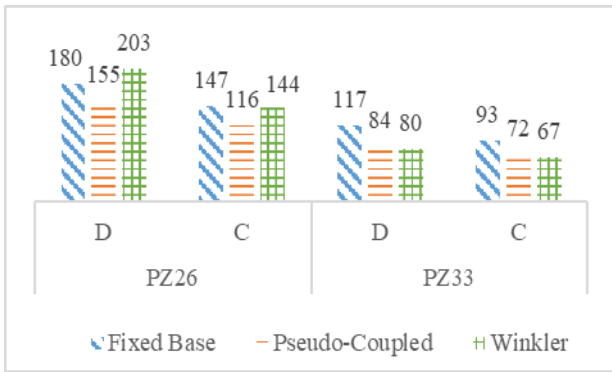


Figure 11. Ey earthquake load of PZ26 and PZ33 shear walls (kN)

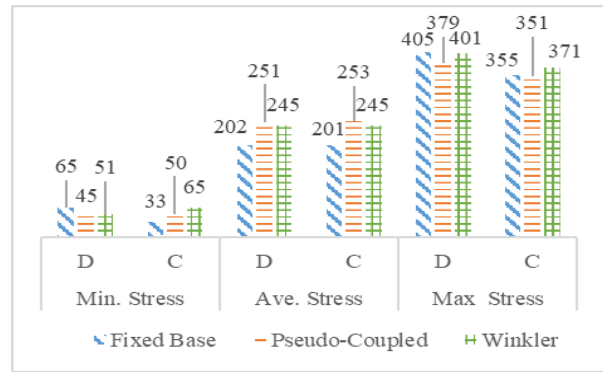


Figure 13. The soil stresses of the raft foundation (kN/m²)

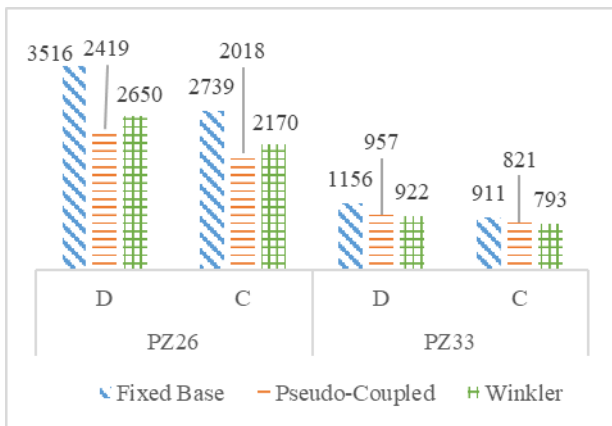


Figure 12. M3 bending moments of PZ26 and PZ33 shear walls (kNm)

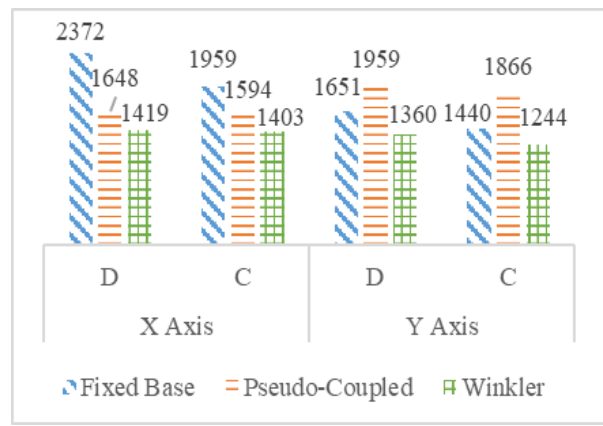


Figure 14. Maximum bending moment of the raft foundation (kNm)

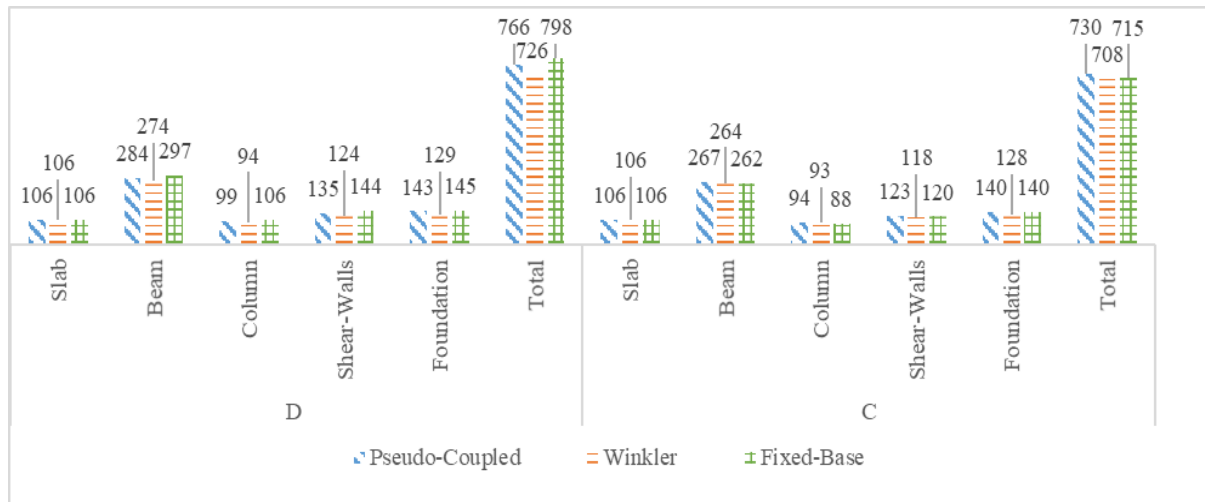


Figure 15. Reinforcement bar quantities obtained from three different analysis results (ton)

4. Conclusions

This study was conducted to examine the effects of different foundation analysis methods on the results of structure analysis. Fixed base and SSI foundation analysis methods are used in 20 storey analysis model. This study also used two different soil groups. The research results in this article have been analyzed under several headings.

- As a result of all the performed analysis, the structure periods of Winkler and pseudo-coupled showed an increase of 30-35% compared to the fixed based model.

- When the horizontal loads and the torsional moments applied on the structures were examined, it was observed that SSI solution was subjected to torsional moments with less horizontal loads of up to 20% compared to fixed base in soil group D.
- Another aim of this study was to examine the effect of the foundation analysis system on the construction economy. The results of the research showed that the most steel requirement has emerged in the fixed base model at soil group D but when analysis models at soil

group C are examined, it has been seen that the most steel requirement has emerged in the pseudo-coupled model.

- It cannot be said that the fixed base analysis system is more secure than SSI solution. Especially when the columns' moment capacities were compared, SSI solutions were much closer to the structural safety limit values
- When analysis results of the shear walls are examined very serious differences were found in shear force and bending moments. Therefore, the buildings analyzed with the fixed base system cannot be said to be on the safer side than the SSI solutions.
- The fixed base foundation system's bending moments give higher results in all soil groups. Therefore, although the fixed base solution remained on the safer side compared to SSI solutions in terms of bending moment, it is still on the more insecure side when examined in terms of the average soil stresses.
- All the results of the analysis showed that the SSI systems give up to 10% more economical results than the fixed base systems in poor soil conditions.

Declaration

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article. The author(s) also declared that this article is original, was prepared in accordance with international publication and research ethics, and ethical committee permission or any special permission is not required.

Author Contributions

All authors contributed to data collection phase. Ö. Özer performed the analysis and wrote the manuscript. S.B. Yüksel made proofreading of manuscript.

References

1. Özer, Ö. and S.B. Yüksel, *Comparing Analysis Results Of Tall Buildings Which Takes Earthquake Effects By Reinforced Concrete Frames And Shear Walls Together According To Tbec 2018 And Tec 2007*. Nigde Omer Halisdemir University Journal of Engineering Sciences, 2020. **9**: p. 931 - 945.
2. Winkler, E., *Die Lehre Von Elasticitaet Und Festigkeit. 1st Edn.* H. Dominicus. 1867, Prague.
3. Vallabhan, C.V.G. and A.T. Daloglu, *Consistent FEM-Vlasov model for plates on layered soil*. Journal of Structural Engineering-Asce, 1999. **125**(1): p. 108-113.
4. Vlasov, V.Z. and N.N. Leontiev, *Beams, Plates And Shells Of Elastic Foundations*. Israel Program for Scientific Translations, Jerusalem, 1966.
5. Girgin, S.C., et al., *The Role Of Soil-Structure Interaction On Structural Design*. Deu Faculty of Engineering Journal of Science and Engineering, 2008. **10**(1): p. 27-37.
6. Daloglu, A.T. and C.V.G. Vallabhan, *Values of k for slab on Winkler foundation*. Journal of Geotechnical and Geoenvironmental Engineering, 2000. **126**(5): p. 463-471.
7. TEC, *Turkish Earthquake Code*. 2007, Ankara: Ministry of Public Works and Settlements.
8. Fatahi, B., S.H.R. Tabatabaiefar, and B. Samali, *Soil-structure interaction vs Site effect for seismic design of tall buildings on soft soil*. Geomechanics and Engineering, 2014. **6**(3): p. 293-320.
9. Tomeo, R., et al., *SSI effects on seismic demand of reinforced concrete moment resisting frames*. Engineering Structures, 2018. **173**: p. 559-572.
10. Moghaddasi, M., et al., *Effects of soil-foundation-structure interaction on seismic structural response via robust Monte Carlo simulation*. Engineering Structures, 2011. **33**(4): p. 1338-1347.
11. Celebi, E., F. Goktepe, and N. Karahan, *Non-linear finite element analysis for prediction of seismic response of buildings considering soil-structure interaction*. Natural Hazards and Earth System Sciences, 2012. **12**(11): p. 3495-3505.
12. Özmen, B., *17 August 1999 Izmit Bay Earthquake Damage Status (With Numerical Data)*, in *TDV/DR 010-53*. 2000: Turkish Earthquake Foundation. p. 132.
13. TBEC, *Turkish Building Earthquake Code*. Ministry of Interior Disaster and Emergency Management Authority. 2018, Ankara.
14. Mısır, S.İ., G. Özden, and S. Kahraman, *The Effects of Using Constant and Variable Subgrade Modulus on the Structural Behaviour*. Sixth National Conference on Earthquake Engineering, 2007: p. 217-228.
15. SCIA, *Foundations and Subsoil (manual)*. 2012, Nemetschek.
16. Farghaly, A.A. and H.H. Ahmed, *Contribution of soil-structure interaction to seismic response of buildings*. Ksce Journal of Civil Engineering, 2013. **17**(5): p. 959-971.
17. Balkaya, C., S.B. Yüksel, and O. Derinoz, *Soil-structure interaction effects on the fundamental periods of the shear-wall dominant buildings*. Structural Design of Tall and Special Buildings, 2012. **21**(6): p. 416-430.
18. Xiong, W., L.Z. Jiang, and Y.Z. Li, *Influence of soil-structure interaction (structure-to-soil relative stiffness and mass ratio) on the fundamental period of buildings: experimental observation and analytical verification*. Bulletin of Earthquake Engineering, 2016. **14**(1): p. 139-160.



Review Article

Comparison of conventional high speed railway, maglev and hyperloop transportation systems

Mehmet Nedim Yavuz ^{a,*}  and Zübeyde Öztürk ^b 

^aMaltepe University, Faculty of Engineering and Natural Sciences, İstanbul and 34000, Turkey

^bİstanbul Technical University, Faculty of Civil Engineering, İstanbul and 34000, Turkey

ARTICLE INFO

Article history:

Received 16 September 2020

Revised 19 December 2020

Accepted 29 December 2020

Keywords:

Conventional high-speed
railway
Hyperloop
Maglev

ABSTRACT

Increasing the speed of transportation has been a subject that human beings have been working on for many years. Because of insufficient traffic corridors, the interest of more passenger in limited time and the advancement of railway technology, the high-speed ground transportation systems have developed. Nowadays, these time-saving transportation systems are becoming important increasingly and systems that will ensure these are implemented and new ones are being researched. High speed ground transportation systems can be divided into three categories: Conventional high-speed railway, Maglev and Hyperloop transportation system. Within the scope of this study, it is aimed to investigate the advantages and disadvantages aspects of high speed transportation systems by comparing in terms of speed, capacity, energy consumption, cost and environmental effects. Conventional high-speed railways can provide high capacity, comfort and safety and reliability thanks to enormous operational experience as regards other systems. The Maglev technology can offer more remarkable travel times, energy efficient and better operational performance. However, high investment cost and incompatibility with other modes are seen the disadvantage features of this technology. The Hyperloop technology is a considerable innovative transportation system which is popularized with publishing design document by Elon Musk in 2013. Projected high-speed and appealing travel times can be evaluated as the advantages of the Hyperloop transportation system. However, there are safety, reliability, comfort and engineering design challenges to overcome in this technology. Taking everything into consideration, the Hyperloop transportation system has potential to be an alternative mode to other systems.

© 2021, Advanced Researches and Engineering Journal (IAREJ) and the Author(s).

1. Introduction

Increasing the speed of transportation has taken an important place among the efforts of human being for a long time. Because of incremental environmental impacts such as traffic jams, prolonged travel times, air pollution and noise caused by existing modes of transportation, countries have changed transport policies. As a result of inadequate traffic corridors, the demand of more passenger in shorter time and the evolution of the railway technology, the high-speed ground transportation systems have developed [1]. High-speed ground transportation systems can be divided into three categories based on the type of technology which is used. These are Conventional high speed railway, Maglev and Hyperloop transportation

system.

In Conventional high-speed railway transportation systems, steel-rail and steel-wheel technology is used. Conventional high-speed railways can be defined in terms of infrastructure, rolling stock and operating conditions. With respect to minimum speed limitations, it is described that a high-speed rail line allows to operate at speeds of over 250 km/h for newly constructed lines or over 200 km/h on existing lines [2].

Maglev technology is described as an abbreviation of the word “magnetic levitation” which is a transportation system comprises of a vehicle that is lifted and pushing by means of magnetic forces along a guideway without physical contact. The main idea behind of Maglev

* Corresponding author. Tel.: +90 216 626 10 50 / 2495.

E-mail addresses: mehmetmedimyavuz@maltepe.edu.tr (M. N. Yavuz), ozturkzu@itu.edu.tr (Z. Öztürk)

ORCID: 0000-0001-9571-9146 (M. N. Yavuz), 0000-0002-2962-6459 (Z. Öztürk)

DOI: 10.35860/iarej.795779

This article is licensed under the CC BY-NC 4.0 International License (<https://creativecommons.org/licenses/by-nc/4.0/>).

technology is that the identical pole of magnets repels while opposite pole of magnets attracts each other. The working principle of the Maglev train is based on three basic principles. These are magnetic levitation, lateral guiding and propulsion principle.

Although the Hyperloop transportation framework was brought to public attention in 2013 by Elon Musk as a fifth mode of transportation after car, airplane, train and ship, the idea of this technology dates back to 1900s. In 1910s, the American rocket pioneer Robert Goddard proposed a floating train was named as vacetrain (vacuum train) inside a vacuum-sealed tunnel from Boston to New York [3,4]. The idea was popularized by Elon Musk with introducing Hyperloop Alpha concept. The main philosophy of Hyperloop technology is that movement of the capsules in lower pressurized tube by means of linear asynchronous motors fed by solar energy. The goal is to reach supersonic speed (1220 km/h) by minimizing air resistance and friction [5].

There are many previous studies on these types of transportation system. While most of these studies compare Maglev and Conventional high speed railway from different technical aspects, few studies aim to contrast these three types of transportation system together. In the study of Liu and Deng [1], the aim is to compare of Maglev and Conventional high speed railway system from operating perspective given corridor from Beijing to Shanghai. As a result of this study, the positive and negative sides of both transportation systems are overviewed in order to select the proper technology for giving corridor.

In the master thesis [6], Maglev and Conventional high speed railway systems are analyzed and contrasted to improve current transportation networks in the United States of America. The results indicate that Maglev technology has more positive sides than Conventional high speed system. In the other master thesis [7], three high speed transportation systems are evaluated in terms of energy consumption and sustainability. The simulation results show that Hyperloop system is best high-speed transportation system with regard to consuming and regenerating of energy.

In the study of Çodur [8], features of Maglev technology and its applications in terms of cost around the world are included. As a result of this study, it is concluded that at which distances Maglev technology is more feasible.

The study [9], comprises of features and history of three transportation systems. It evaluates challenges of application about three systems.

Janić [10] evaluates the three transportation systems, assuming corridor between Moscow-St. Petersburg. It is indicated that Hyperloop system can be an alternative to Transrapid Maglev and Conventional high speed rail.

Contreras [11] reviews of features of three systems and emphasizes current and future challenges.

The study of Armağan [12] presents a review of the Hyperloop transportation system. In the study, safety, resistance of weather and earthquake are remarked as the advantage side of Hyperloop. However, in terms of the cost issue, this technology is remarked as expensive.

In the scope of this study, it is aimed to reveal advantage and disadvantage aspects of three transportation systems by comparing in terms of speed, capacity, energy consumption, cost and environmental impacts. The contribution of this paper is to highlight benefits and challenges of these three high speed transportation systems by presenting state of art review.

2. Conventional High Speed Railways

Due to rapid urbanization process across the globe, congestion and inadequate capacity in existing highway and airway modes of transportation, have enabled Conventional high speed railways to be promoted in transport policies. Moreover, reduced travel times thanks to high speed offered by Conventional high speed railways, as well as providing a safe and comfortable trip have increased the popularity of high speed railway system [1].

It is not possible that define Conventional high speed railway based on only one factor. The high-speed railway technology is a complex system that consists of many components such as infrastructure, rolling stock, telecommunication, operating conditions and equipment etc. According to European Union Directive 96/48/EC Annex 1, in terms of infrastructure, it is defined as building track specially or upgraded for high speed travel. In terms of minimum speed limit, it is defined as minimum speed of 250 km/h for newly constructed lines and about 200 km/h on existing lines which have been particularly promoted. This must apply to at any rate segment of the track. Regarding to operating conditions, rolling stock needs to be designed along its infrastructure for whole compatibility [2].

The history of conventional high speed railway traced back to 1964 with the opening of Tokaido Shinkansen in Japan. After Shinkansen enormous success, many European countries, especially France, Germany, Italy developed new technologies in order to increase share of railway in the transportation sector. France in 1981, Italy and Germany in 1988, Spain in 1992, China in 2003, Turkey and Netherland in 2009 opened first high-speed lines all around world. As a result of huge investment by China, this country has a largest conventional railway network in today [2].

Nowadays, there are tens of thousand kilometer lines in operation or plan from different countries around the

world. Table 1 presents that the total current length of high-speed lines of some countries according to International Union of Railway (UIC) data in 2020.

3. Maglev Technology

The Maglev term is an abbreviation of “magnetic levitation”. Magnetic levitation is defined as a transportation system consisting of a vehicle that is lifted and pushed with help of magnetic forces along a guideway without physical contact.

The main working principle of Maglev trains is based on repelling of equal polar magnets and attraction of opposite polar magnets each other. Magnetic levitation, lateral guiding and propulsion functions are provided by means of magnetic force. The force generated by electromagnets creates a distance of approximately 1 centimeter between the guideway and bottom of the train. The distance between train and guideway is controlled precisely. Owing to lateral guiding, the lateral stability and straddling of the train are ensured. The forward motion and braking of train are same principle as the electric motor. Motor windings lined along guideway acts as a stator and electromagnets on the train acts as a rotor. The alternating current supplied to motor windings creates a magnetic field that moves the train.

The speed of Maglev trains can be adjusted depending on the frequency of the alternating current supplied. Braking and stopping of the train are performed by generating a magnetic force in the opposite direction [14, 15].

There are two major types of Maglev technology. These are Electromagnetic Suspension System (EMS) and Electrodynamic Suspension System (EDS). Fundamentally, while EMS uses magnetic attraction force to move in the air, magnetic repulsion force is used in EDS system (Figure 1). Table 2 summarizes that difference of Maglev technology types.

The history of Maglev train can be dated from 1934 with the patent of Hermann Kemper from Germany. In the past few decades since then, as a result of human endeavor’s, first commercial Maglev line is opened in Shanghai [18]. There are six commercial Maglev lines around world by the year of 2020 (Table 3) [17].

Table 1. The lengths of conventional high speed lines of some countries from around world [13]

Country	In Operation (km)	Under Const.(km)	Total (km)
China	35,388	5,250	40,638
Japan	3,041	402	3,443
France	2,734	-	2,734
Germany	1,571	147	1,718
Italy	921	327	1,248
Spain	3,330	1,293	4,623
Turkey	594	1,652	2,246
USA	735	763	1,528

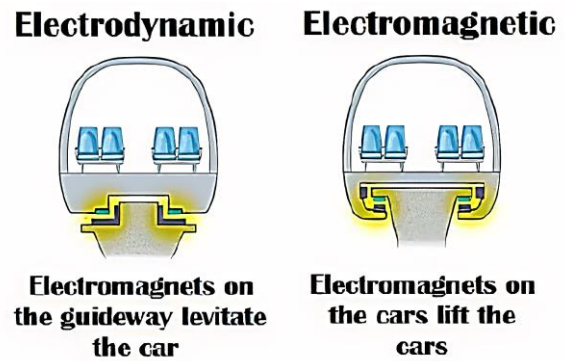


Figure 1. Comparison of EMS and EDS [16]

Table 2. Comparison of maglev technology types [16]

EMS	EDS
Using of conventional magnets	Using of superconducting magnets
About 15 mm air gap between train and guideway	About 15 cm air gap between train and guideway
Need precise control systems	Less sensitive to earthquake
Work at all speeds	Only work at minimum speed about 30 km/h
Less energy consumption	Reaching higher speeds

Table 3. Existing maglev lines around world [17]

Country/City	Line	Opening	Length (km)
China, Shanghai	Airport Maglev	2004	30.5
Japan, Nagoya	Linimo Metro	2005	8.9
South Korea, Daejeon	Museum Maglev	2008	1
South Korea, Incheon	Airport Maglev	2016	6.1
China, Changsha	Airport Maglev	2016	18.6
China, Beijing	S1 Metro	2017	10.2

4. Hyperloop

The Hyperloop technology was announced in a design document called Hyperloop Alpha as a new high-speed transportation concept in 2013 by Elon Musk, CEO of SpaceX and Tesla companies. This new system is based on the movement of the capsule with the help of linear electromagnetic motors through tube with low air pressure at similar or higher speeds than air travel. It is stated that Hyperloop can travel between Los Angeles and San Francisco in 35 minutes [5].

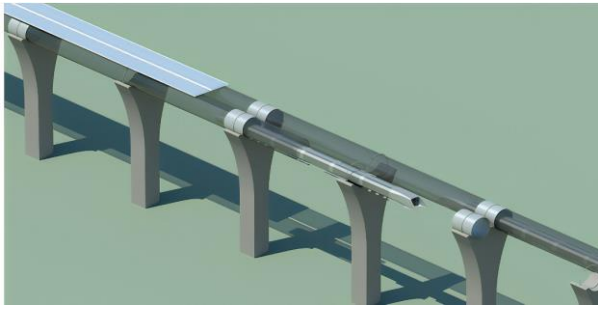


Figure 2. Hyperloop vacuum tube design [3]

The main working principle of this emerging technology is to minimize air friction in a vacuum tube with magnetically levitated capsule. The Hyperloop technology consists of vehicle (capsule) which is used for passenger or cargo transportation, vacuum tube and propulsion system. The capsule is the main component of Hyperloop technology responsible for passenger or cargo transport.

Due to the targeted high speed, use of wheels is not possible in capsule design. The shape of capsule is chosen so that there is a minimal resistance to air friction during movement. The Hyperloop route consists of a cylinder vacuum tube. It is proposed to reduce the cost of construction and keep the required construction area size to a minimum, vacuum tube is built on columns that are constructed at intervals (Figure 2). In order to speed up and slow down the capsule, the linear asynchronous motor would be used. The stationary motor element (stator) would be constructed at various positions along the length of the tube to accelerate the capsule, while the movable motor element (rotor) would be placed in capsules to transfer momentum to the capsule through linear accelerators [5].

The Hyperloop Alpha design document which was announced by Elon Musk in 2013, attracted great attention worldwide. After releasing of design document, many companies and academic teams were established to work on the Hyperloop technology. Some of these prominent companies; Hyperloop One, Hyperloop Transportation Technologies, Delft Hyperloop, Hardt Hyperloop etc. In this context, the first test track called DevLoop was built by Hyperloop One company in the Nevada desert in the north of Las Vegas. This test track is approximately 500 meters long, and it is used to test prototypes of passenger and cargo capsules [19]. While the United States mostly leads work on the Hyperloop technology, countries such as the United Arab Emirates, India, France and Netherlands are conducting research and studies on the application of this system to their own countries.

5. Comparison of Transportation Systems

It is possible to compare Conventional high speed railway, Maglev and Hyperloop transportation systems from different technical aspects. Comparison of these

three high speed transportation systems can be categorized as listed below;

- General comparison of three transportation systems in terms of speed, capacity, compatibility of system etc.
- Comparison in terms of geometric standards
- Comparison in terms of cost
- Comparison in terms of energy consumption
- Comparison in terms of environmental effects

5.1 General Comparison of Transportation Systems

Travel time is one of the most important parameters that cause passengers to choose one mode of transportation instead of another. The design and operating speed of each mode of transportation play vital role in terms of travel time. While design speed is generally dependent on infrastructure, maximum operating speed varies depending on technical characteristics and operating models of the train. The average speed is defined as track length divided by total travel time.

Table 4 shows that a comparison of the design speed, operating speed and average speed of Shanghai Maglev, TGV, ICE high speed railway trains and Hyperloop. For the Hyperloop transportation system, speed values which are stipulated in the Alpha report were used. It is desired to make comparison based on Table 4, Maglev technology has advantage over Conventional high speed railways in terms of maximum operating speed. Although there is not a commercially operated line with Hyperloop, development studies on speed continue. Currently, the highest speed pod which is nearly 463 km/h, is acquired by a team from Technical University of Munich during the Hyperloop pod competition [20]. If desired speeds would be reached, the Hyperloop transportation system could be the most advantageous type of transportation.

Capacity is a parameter that affects the choice of types of transportation system by passenger. The capacity of a railway line can be determined by factors such as the number of passenger per train section, the number of train section and headway. While these factors mostly depend on technical characteristics, concepts such as population density and passenger demand, have a role in determining the capacity of the line. When comparing among Conventional high speed trains with regarding capacity, Shinkansen trains stand out as the highest capacity train. While the Shanghai Maglev train has a capacity of 574 passengers, 28 passengers per capsule is proposed in the Hyperloop transportation system. [5, 21].

Table 5 shows that comparison of evaluated transportation systems in terms of capacity and headway. The value of capacity in the table, is obtained as maximum passenger per hour in one direction.

Table 4. Comparison in terms of speed [1, 5]

Speed (km/h)	TGV	ICE	Shanghai Maglev	Hyperloop
Design Speed	350	330	550	1220 (theoretical)
Operating Speed	300	280	430	Not Available
Average Speed	250	200	290	965 (theoretical)

Table 5. Comparison in terms of capacity, headway [1, 5]

Type	TGV-D	SKS-E4	Shanghai Maglev	Hyper loop
Number of Section	12	16	6	1 capsule
Seat Capacity	1090	1634	574	28
Headway (min)	5	3	15	2
Capacity	13080	32680	2296	840

It is observed that Conventional high speed trains are advantageous in terms of capacity based on data in Table 5.

When the three modes of transportation are evaluated in terms of the compatibility of the system, Conventional high-speed railways have a superior advantage over the Maglev and Hyperloop transportation systems, thanks to offering the opportunity to use existing infrastructure and thus connect with existing rail networks. Since Hyperloop and Maglev transportation systems have their own special tracks, it is not possible to work with other available modes. This feature can be shown as one of the disadvantages of Hyperloop and Maglev transportation systems.

Conventional high speed railway is a type of transportation that has proven itself in terms of safety and reliability, along with lines operating in many countries around the world. To illustrate, Shinkansen trains from Japan, TGV trains from France and ICE trains from Germany have been serving passengers safely and comfortably for years. In the Maglev transportation system, it is designed to largely eliminate security risks. Examples of these design patterns are the train's motion with wrapping guide-way and precise control of the distance between the train and the guide-way. No fatal accident has occurred since the Shanghai Maglev line was commissioned. In the Hyperloop transportation system, the design is planned to be done so that the tube and capsule are not affected by natural disasters such as earthquakes. When evaluating in terms of passenger comfort, safety and reliability among three modes of transportation, Conventional high-speed railways and the

Maglev transportation system offer a safe and comfortable trip. In the Hyperloop concept, although there are concerns in terms of comfort at the targeted sound levels, there is no data that can be compared since there is no commercially operated Hyperloop line for now.

5.2 Comparison in terms of Geometric Standards

Track geometry is a very important factor for train behavior. The notions related to track geometry can be sorted as track gauge, cant, transition curve, horizontal and vertical curve radius, longitudinal grade. In case of comparing with regarding geometric standards, Maglev and Hyperloop have advantages such as moving at the lower curve radius with the same speed and travelling at higher speeds in the same curve radius, and climbing higher slopes compared to Conventional high speed railways (Table 6) [22].

In the Hyperloop Alpha study report, in order to control g-force which arises from acceleration-deceleration and change of direction, the velocity plan and optimum horizontal curve radius were stated in a table. This table is obtained based on minimizing the effects of g-force, curvature radius and speeds specific to terrain between San Francisco and Los Angeles and the maximum 0.5 g force on passengers. These values are given in Table 7.

In the comparison of cant, the maximum allowable value in the Conventional high speed railways is taken as 180 mm. In the Maglev transportation system, cant is not represented in mm. The maximum allowable cant is 12° equivalent to approximately 310 mm. In some exceptional cases, it can be increased to 16°, that is approximately 410 mm. This advantage enables Maglev to increase alignment flexibility according to Conventional high speed railway [23]. In the Hyperloop Alpha report, there is not information about cant. It is a one of the uncertain issues for Hyperloop technology.

Table 8 compares the minimum vertical curve radius that Conventional high speed railways and Maglev transportation systems should provide at given speeds.

Table 6. Comparison in terms of horizontal curve radius (m) [22]

Design Speed (km/h)	ICE-03	Transrapid Maglev
Max. Side Acceleration	1 m/s ²	1,5 m/s ²
200	1400	705
250	2250	1100
300	3200	1590
350	-	2160
400	-	2825
450	-	3580

Table 7. Minimum horizontal curve radius in Hyperloop [5]

Track Route	Design Speed (km/h)	Min. horizontal curve radius (m)
L.Angeles-Grapevine South	480	3670
L.Angeles-Grapevine North	890	12550
I-580/San Francisco Bay	1220	23500

Table 8. Comparison in terms of vertical curve radius (m) [22]

Design Speed (km/h)	ICE-03		Transrapid Maglev	
	Crest	Sag	Crest	Sag
Max. Vertical Acceleration	0,5 m/s ²	0,6 m/s ²	0,6 m/s ²	1,2 m/s ²
200	6400	5200	5150	2600
300	14400	11700	11600	5790
330	17400	14200	14000	7000
400	-	-	20600	10300
450	-	-	26000	13000

For the Hyperloop technology, the vertical curve radius is not explicitly given in the Alpha design document. However, in order to diminish earthworks, pipelines are constructed as compatible with road geometry. Therefore, the grade of vertical curve will not exceed %6 according to the AASHTO highway specification.

5.3 Comparison in terms of Cost

Cost plays an important role in selecting and evaluating any mode of transportation. There are different cost concepts such as construction, operating and maintaining cost for a mode of transport. The cost of construction includes the cost of constructing the line and stations, the train control system, and the purchase of the trains. The operating cost includes the expenses required to keep the line operating. Maintenance costs include the necessary expenses to operate the line properly and efficiently.

It is seen that the construction cost of Conventional high speed railways varies around the world. The total construction cost of a mode of transport depends on the chosen technology, design speed, land topography, land acquisition requirements, the need for special structures and auxiliary facilities. To illustrate, in the study conducted on 45 projects, it is seen that the construction costs of Conventional high speed railways vary between 6 and 45 million Euros per kilometer. The average cost was found to be 17.5 million Euros. It is occurred that Conventional high speed railway construction costs in the Asian continent are higher than in the European continent. For the European continent, the construction costs in France and Spain are slightly lower than in Germany and Italy. Although the

topography of land has an effect on this difference, the construction method also emerges as an important factor [24].

Maglev transport type has higher construction costs than Conventional high-speed rail due to the completely separate right of way, construction of special facilities and incompatible with existing transportation systems.

The fact that Maglev trains cannot run on normal train tracks that they are not compatible with the existing railway infrastructure that they require a completely new line are important factors that increase the construction cost.

In the Hyperloop technology, the construction cost covers the construction of the guiding line, capsule production, station construction costs. The Hyperloop Alpha report foresees a construction cost of approximately \$ 6 billion for the passenger-carrying version alone and \$ 7.5 billion for the passenger + vehicle version [5]. It is not included earthwork and special station costs in this cost estimation. Additionally, a current project between Dubai and Abu Dhabi, which is planned to be completed soon, it is estimated to cost between \$ 20 million to \$ 40 million per kilometer [25].

When comparing the average construction costs between these transportation systems, it is seen that the highest construction costs per kilometer belong to the Maglev mode of transport. According to stipulated costs in the Hyperloop Alpha report, the Hyperloop transportation system has a lower construction cost than Maglev. After construction of the track, operation and maintenance costs come into play. Maglev train can be operated at very high speeds without any deterioration and therefore it is more economical to operate than Conventional high-speed railways that require regular maintenance. Transrapid Maglev has fully automatic operation and is driverless. For this reason, personnel costs which are a component of operating costs, are expected to be lower than Conventional high speed railways. Another positive feature of the Maglev transportation system is that maintenance costs are lower than Conventional high speed railways.

Table 9. Comparison in terms of construction cost [5,22]

Route	Transport Mode	Cost (billion \$)	Length (km)	Cost per km (million \$)
Tokyo-Osaka	HSR	0,92	570	1,6
Paris-Lyon	HSR	2,06	1000	2,06
Madrid-Barcelona	HSR	10,62	620	17,12
Beijing-Shanghai	HSR	35,80	1432	25
Shanghai	Maglev	1,58	30	52,67
Linimo	Maglev	0,92	8,8	104,77
L.Angeles-California	Hyper loop	6	563	10,66

Since Maglev trains move in suspension along the guideway, wear and tear damage caused by wheels in high speed railways are eliminated. Since the acceleration and deceleration of Maglev trains are provided by the magnetic force created by electromagnets, there is no contact with the rail.

Since there is no friction between the rail and the train, the maintenance cost of the rail is very low and frequent repairs and controls are not required.

In evaluating the Hyperloop operating cost, the energy consumption cost is expected to be low as the energy requirement will be met by solar panels installed on tubes. Due to the fact that the system has a fully automatic operation, personnel cost can be considered as minimum. In the Hyperloop transportation system, the maintenance cost is expected to be low due to the lack of mechanical friction and weather protection as in the Maglev system. Although there is not operating experience of Hyperloop technology, the operating plans are expected to similar by airplanes. However, the security checks and terminal waiting times will be shorter and maintenance cost will be less than aircraft.

From Figure 3, comparing the Germany ICE high-speed train and the Maglev Transrapid train in terms of maintenance costs; the Maglev transportation system has about 59% less regarding to the maintenance cost of the train, 71% less of the maintenance cost of the guideway and totally 66% less maintenance cost than the Conventional high-speed railway.

In the context of ticket fee, it is indicated as approximately \$ 20 for one trip according to Elon Musk’s math in the design document. Moreover, a study which is about the proposed route between Pittsburg and Chicago, ticket fee is stated approximately \$60 [27]. However, the single ticket cost of Shanghai Maglev is 50 yuan, which equals approximately \$8[28]. In the light of this information, it can be commented that ticket cost of Hyperloop will be higher than Maglev.

5.4 Comparison in terms of Energy Consumption

The transportation sector is an important source of greenhouse gas emissions. For this reason, new technologies that reduce energy consumption in the transport sector are supported. Examples of these technologies are weight reduction (producing lighter trains, etc.), energy efficient driving techniques, reducing aerodynamic friction, and regenerative braking system. Electricity is used as an energy source in both conventional high-speed railways and Maglev.

In the Hyperloop technology, it is planned to meet the energy need with solar panels built on tubes. The Maglev train is much more economical compared to Conventional high speed railways, without contact with the guideway, high efficiency linear motor and low aerodynamic

resistance. In general, the Maglev transportation system consumes 20-30% less energy than Conventional high speed railways. Convenient aerodynamic features and non-contact technology make Maglev cost-effective about energy consumption. In the Hyperloop Alpha study report, it is emphasized that Hyperloop technology will operate entirely with solar energy. Therefore, it is expected to have less energy consumption than Maglev and Conventional high speed railways. However, since there is no operational experience in the Hyperloop transportation system yet, it is not known exactly how much the energy consumption will be. In the figure 4, it is given comparison of Conventional high speed railways and Maglev trains in terms of energy consumption at certain speeds.

When the energy consumption values of Conventional high speed railways and Maglev trains are compared at 330 km / h, the Transrapid Maglev train has the lowest energy consumption with 45 Watt hours / seat / km, while the ICE-3 train has 59 Watt-hour / seat / km has the highest energy consumption. As a result, it is seen that Maglev train has approximately 31% less energy consumption than Conventional high speed railways (Figure 4).

5.5 Comparison in terms of Environmental Effects

In recent years, concerns about environmental effects of the transportation sector have been increasing. Due to being environmentally friendly of the rail transport system, it has an advantage over other transport modes.

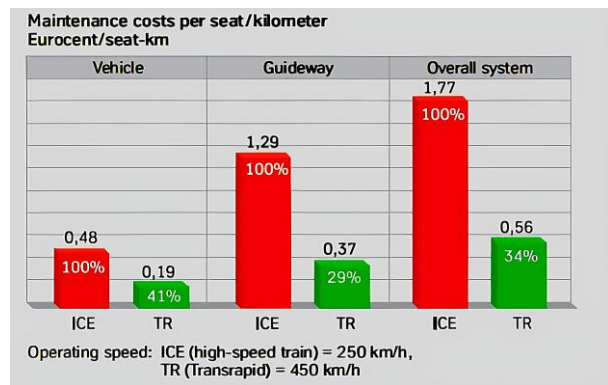


Figure 3. Comparison of maintenance cost [26]

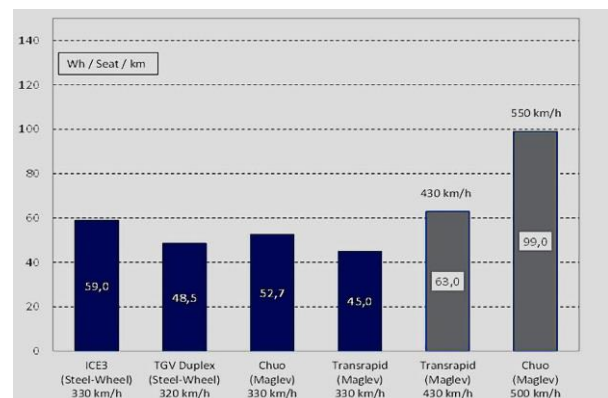


Figure 4. Energy consumption in Wh/Seat/km [29]

It is possible to divide the environmental impacts of Conventional high speed railways, Maglev and Hyperloop transportation types into sub-headings such as land use, noise, vibration.

The amount of land using depends on the type of transport system and the way the guideway is built. At grade tracks with land, consumes land which is below the track as well as catenary and signal poles are located next to it.

However, elevated track only uses the land holding columns supporting the track. The land below elevated track can be used for different purposes. As a result, elevated tracks have advantage in land using compared to track which is at grade with land. In the Maglev transportation type, lines are generally constructed as raised with columns. Therefore, Maglev transportation type has less land use than Conventional high-speed railways. In the Hyperloop concept, although the land use is not known due to lack of operational experience, as the construction method is similar to the Maglev, it is predicted to use less land compared to Conventional high speed railways. The Maglev transportation system has a distinct advantage over Conventional high speed railways in terms of land use in mountainous areas. Hyperloop transportation type is predicted to have land use values close to Maglev transportation type in terms of land use (Table 10).

Noise is a problem not only for passengers, but also for those living near the line corridor. Significant progress has been made in the control and mitigation of transport noise in recent years. Noise emissions can be grouped as noise from propulsion system, mechanical noise from wheel-rail interaction or guideway vibrations, and aerodynamic noise. Aerodynamic noise prevails at higher speeds, while at lower speeds (speeds below about 200 km/h) noise from the propulsion system or machinery is dominant. In the Maglev system, mechanical noise is not observed at low speeds since physical contact is eliminated with non-contact technology.

When the noise levels are compared at different speeds, it appears that Maglev technology is quieter than the Conventional high speed railway at speeds of 200-300 km/h. In the Hyperloop concept, it is prevented noise from moving the capsule thanks to low air pressure inside the tube. The sole possible source of noise is caused by vacuum pumps however; it is assumed to be minor [31].

Table 10. Comparison of land using [30]

Type	HSR	Maglev	Hyper loop
Average Land Consume (m ² /m)	26,2 (Plain) 43,5 (Hilly)	23,2 (Plain) 24,4 (Hilly)	Similar with Maglev

Thanks to non-contact technology in the Maglev transportation system, the Maglev train causes less vibration than Conventional high-speed trains. In the Hyperloop transportation system, as the capsule is aimed to move in the air inside the tube, it is predicted that the vibration level would be lower than Conventional high-speed trains.

6. Results and Discussion

These three high speed transportation systems are compared with regarding different factors such as speed, capacity, compatibility of the system, geometric standards cost, energy consumption and environmental effects in Table 11.

When it is made an evaluation based on parameters that are compared between examined transportation types; Hyperloop and Maglev transportation system have advantage over Conventional high speed railways in terms of speed, travel time, acceleration and braking rates. To illustrate, the route between San Francisco and Los Angeles takes 130 minutes with California high speed train, 116 minutes by Maglev train and 35 minutes by Hyperloop capsule. If target speed values for the Hyperloop concept could be reached, the Hyperloop technology would have a serious advantage in speed and travel time parameters.

If it is evaluated in terms of capacity and headway parameters, it is seen that Conventional high speed railways have a significant advantage over Maglev and Hyperloop. For instances, in the light of the data in Table 5, the maximum passenger capacity value per hour in one direction for Conventional high speed railways is approximately 14 times of the Maglev and approximately 38 times of the Hyperloop.

In terms of compatibility of the system, Maglev and Hyperloop transportation system are not compatible with existing lines because it requires its own special line structure. The Conventional high speed railway has advantage over other types of transportation thanks to their compatibility with existing lines.

When it is analyzed in terms of safety, reliability and comfort, Conventional high-speed railways offer comfortable and safe travel to millions of passenger in many countries since the first high-speed rail line in Japan was commissioned in 1964. In the Maglev transportation type, no accident was reported on the Shanghai Maglev line since 2004 and other urban lines operated in other countries.

In the matter of geometric standards, Maglev and Hyperloop have advantages such as moving at lower radius of curves at the same speed compared to Conventional high speed railways, and being able to travel at higher speeds in the same curve radius, and to climb higher slopes. Likewise, the maximum allowed cant value is higher.

Table 11. Summarizing comparison parameters

System Features	Conventional HSR	Maglev	Hyperloop (HL)
Maximum Speed (km/h)	241 Acela (Boston to Newyork) 270 TGV (Paris to Lyon)	430 (Shanghai)	1220 (theoretical)
Capacity	1000 per train (California HSR)	574 per Shanghai Maglev train	28 per capsule
Use of existing infrastructure	New lines combined with existing lines	Need special track	Need special track
Construction Cost	Lower than Maglev and higher than HL (likely)	Higher than HSR	Lower than HSR (likely)
Operating and Maintenance Cost	Higher than Maglev and HL	Lower than HSR	Lower than HSR (likely)
Energy Consumption	Higher than Maglev and HL	Less than HSR	Solar panels Less than HSR (likely)
Safety	Proven technology	Proven technology	At concept stage
Noise and Vibration	Higher than Maglev and HL	Less than HSR	Less than HSR (likely)
Land Consume	Higher than Maglev and HL	Less than HSR	Less than HSR

It is seen that highest construction cost belongs to the Maglev transportation type. Incompatibility with existing lines and high construction costs can be considered as negative aspects of the Maglev transportation system. In the Hyperloop Alpha report, the total construction cost of the 563-kilometer route between San Francisco and Los Angeles is stated as \$ 6 billion. It should be noted that earthworks and construction of special station are not included in this cost estimation. Taking this average cost into consideration, the Hyperloop transportation system is cheaper than the Conventional high speed railway in terms of construction cost. However, cost issue is one of challenges to overcome for the Hyperloop transportation system.

Since Maglev train and Hyperloop capsule move without contact with guideway, the maintenance cost is less than Conventional high speed railways.

It is made a comparison based on energy consumption, Maglev trains consume 20-30% less energy than Conventional high speed trains thanks to aerodynamic feature of train and contactless technology. In the Hyperloop Alpha report, it is stated that Hyperloop capsule would run entirely on solar energy. As a result, Hyperloop mode of transport would be expected to have less energy consumption than Maglev and Conventional high speed rail.

When evaluating in terms of the land using, Maglev and Hyperloop transportation types are designed as elevated tracks. Thus, savings in the land use are provided.

Table 12. Comparison and classification of references

Major Areas	Economic Analysis	Engineering Design	Environmental Issues
Ref.	[5], [22], [24], [25], [26], [27], [28], [29]	[1], [2], [5], [14],[15],[16] [21],[22],[23]	[5], [29] [30], [31]

Moreover, thanks to the contactless technology used in the Maglev and Hyperloop transportation system, less noise and vibration values are achieved.

Finally, in this section, the used references in this study are categorized as economic analysis, engineering design and environmental issues. The classification and comparison of references are given in a table (Table 12).

7. Conclusions

Scientific and technological research continue throughout the world to develop increasingly faster transportation systems. As a result of these studies, after the Conventional high-speed railways, which are operated in many countries around the world and offer millions of passengers the opportunity of fast and comfortable travel, Maglev technology pioneered by Japan and Germany, and finally a new mode of transportation called as Hyperloop was introduced with Hyperloop Alpha report by Elon Musk in 2013. Within the scope of this study, the advantages and disadvantages of three transportation systems against each other were examined. As a result, although Hyperloop technology is a highly innovative mode of transportation, works on the development of the technology continue.

The projected low construction cost, speed and minimum environmental impacts made it stand out among the transportation types studied.

Reliability and comfort uncertainty due to not having operational experience yet, and the lack of capacity appear to be the negative aspects of this type of transportation. Taking all parameters into consideration, it is demonstrated that Hyperloop transportation system can be an alternative to Conventional high speed railways and Maglev.

Declaration

The authors declared no potential conflicts of interest

with respect to the research, authorship, and/or publication of this article. The authors also declared that this article is original, was prepared in accordance with international publication and research ethics, and ethical committee permission or any special permission is not required.

Author Contributions

All authors reviewed the sources together. M.N. Yavuz wrote the manuscript. Z. Öztürk supervised the study and made proofreading of manuscript.

References

- Liu, R., Deng, Y. *Comparing Operating Characteristics of High-Speed Rail and Maglev Systems: Case Study of Beijing-Shanghai Corridor*. Transportation Research Record, 2004. **1863**(1): p. 19-27.
- UIC High Speed Rail Report. [cited 2020 6 July]; Available from: https://uic.org/IMG/pdf/uic_high_speed_2018_ph08_web.pdf.
- Van Goerverden, K., Milakis, D., Janic, M., & Konings, R. *Analysis and modelling of performances of the HL (Hyperloop) transport system*. European Transport Research Review, 2018. **10**(2): p. 1-17.
- Gieras, J. *Ultra high-speed ground transportation systems: Current status and a vision for the future*. Przegląd Elektrotechniczny, 2020 : p.1-7.
- Tesla, Hyperloop Alpha Report. [cited 2020 20 July]; Available from: https://www.tesla.com/sites/default/files/blog_images/hyperloop-alpha.pdf
- Ziemke, D. *Comparison of High-Speed Rail Systems for the United States*. Msc Thesis, Georgia Institute of Technology, 2010.
- Riviera, M. *High-Speed Trains Comparison to Hyperloop: Energy and Sustainability*. Msc Thesis, Politecnico Di Torino, 2017.
- Çodur, M. *Türkiye’de Maglev Trenlerinin Uygulanabilirliğinin Araştırılması*. Journal of the Institute of Science and Technology, 2017. **7**(1) : p. 207-215.
- Gonzalez-Gonzales, E., Nogues-Linares, S. *Railways of the future: Evolution and Prospects of High-speed, Maglev and Hyperloop (1st Part)*. DYNA, 2017. **92** (4): p. 371-373.
- Janić, M. *Multicriteria Evaluation of the High Speed Rail, Transrapid Maglev and Hyperloop Systems*. Transportation Systems and Technology, 2018. **4**(4): p. 5-31.
- Contreras, M. *HSGT Systems: HSR, Maglev and Hyperloop*. The Journal, 2018.
- Armağan, K. *The fifth mode of transportation: Hyperloop*. Journal of Innovative Transportation, 2020. **1**(1): p. 1105.
- UIC High Speed Lines in the World Report. [cited 2020 6 July]; Available from: https://uic.org/IMG/pdf/20200227_high_speed_lines_in_the_world.pdf
- Yadav, M., Mehta, N., Gupta, Aman., Chaudhary, A., Mahindru, D. *Review of Maglev Levitation (MAGLEV): A Technology to propel Vehicles with Magnets*. Global Journal of Researches in Engineering Mechanical & Mechanics, 2013, **13** (7-A).
- Solak, K. *Raylı Sistemlerin Alternatifleri ile Manyetik Yastık Üzerinde Hareket Eden Trenlerin (Maglev) Çok Ölçütlü Değerlendirme Yöntemi ile Karşılaştırılması*. Msc Thesis, Gazi University, 2011.
- Jahan, F., Parveen, A., Bisht, S. *Feasibility Study of Maglev Trains on Existing Indian Railways Infrastructure*. International Journal on Emerging Technologies, 2014. **5**(2): p. 106-109.
- [cited 2020 10 July]; Available from: <https://www.maglev.net/all-existing-and-under-construction-maglev-lines>.
- Lee, H., Kim, K., Lee, J. *Review of Maglev Train Technologies*. IEEE Transactions on Magnetics, 2006. **42**:p.1917-1925.
- Virgin Hyperloop. [cited 2020 20 July]; Available from: <https://virginhyperloop.com/project/devloop>
- [cited 2020 20 July]; Available from: <https://www.sciencealert.com/there-s-a-new-record-for-travelling-at-the-speed-of-hyperloop>
- MaglevBoard. [cited 2020 10 July]; Available from: <https://www.maglevboard.net/en/facts/systems-overview/transrapid-maglev/transrapid-maglev-shanghai>
- Witt, M., Herzberg, S. *Technical-economical System Comparison of High Speed Railway Systems*. [cited 2020 10 July]; Available from: http://www.maglev.ir/eng/documents/papers/conferences/maglev2004/topic1/IMT_CP_M2004_T1_11.pdf
- [cited 2020 20 July]; Available from: <https://pwayblog.com/2016/09/07/maglev-guideway-design/>
- Campos, J., De Rus, G., Barron, I. *Economic Analysis of High Speed Rail in Europe*. BBVA Foundation, 2009.
- [cited 2020 18 October]; Available from: <https://www.khaleejtimes.com/technology/abu-dhabi-hyperloop-to-cost-up-to-40-million-per-kilometre-->
- Monorails Australia. [cited 2020 20 July]; Available from: <https://www.monorailsaustralia.com.au/Maglev.pdf>
- [cited 2020 20 October]; Available from: <https://www.wpxi.com/news/top-stories/study-how-much-would-hyperloop-ticket-cost-pittsburgh/LOBFEAVJFRF7HL5WMHL6BXOY6Y/>
- [cited 2020 20 October]; Available from: <http://www.smtdc.com/en/jszl.html>
- Fritz, E., Blow, L., Klühspies, J., Kircher, R. *Energy Consumption of Track-Based High-Speed Trains: Maglev Systems in Comparison with Wheel-Rail Systems*. The International Maglev Board, 2018. **4** (3 suppl 1):p.134-155.
- Wang C., Wang K. *A Study on Environmental Impact of High Speed Maglev Traffic Engineering*. International Conference on Management and Service Science, 2010, p.1-3.
- Van Goeverden, K., Milakis, D., Janic, M., Konings, R. *Performances of the HL(Hyperloop) Transport System*. Proceedings of the BIVÉC-GIBET Transport Research Days 2017: Toward an Autonomous and Interconnected Transport Future, 2017, p. 29-43.



Review Article

An analysis of content-based image retrieval

Hakan Koyuncu ^{a,*} , Manish Dixit ^b  and Baki Koyuncu ^a 

^aComputer Engineering Department, Altınbas University, Istanbul, 34217, Turkey

^bDepartment of CSE/IT, MITS, Gwalior (M.P), 517325, India

ARTICLE INFO

Article history:

Received 20 October 2020

Revised 02 February 2021

Accepted 12 February 2021

Keywords:

Content-Based Image Retrieval (CBIR)

Feature Extraction

Performance measure

Shape

Texture

Tamura

Wavelet transform

ABSTRACT

Nowadays, working on digital images is gaining much popularity in multimedia systems, due to the rapid increase in the utilization of large image databases. Thus, the Content-Based Image Retrieval (CBIR) method has become the most valuable method for these databases. This study mainly focuses on content-based image retrieval; which uses image features like color, shape, texture, etc. by searching the user query image from a large image database based on user request. CBIR is the most widely used technique as its searching capability is faster than the other traditional methods, and it works well in retrieving images automatically. It is also a big alternative approach to traditional methods. The CBIR techniques are used in many applications like surveillance detection, crime avoidance, fingerprint identification, E-library, medical, historical monument and biodiversity information systems, and many more. A total of 38 CBIR articles were comparatively analyzed.

© 2021, Advanced Researches and Engineering Journal (IAREJ) and the Author(s).

1. Introduction

Content-based image retrieval (CBIR) also known as Content-Based Visual Information Retrieval (CBVIR) or Query by Image Content (QBIC), is a method for solving image retrieval problems that employs computer vision methods. Content-based image retrieval is conflicting with conventional knowledge of conceptual approaches.

Content-based is searching and analysis of different image features. It is a kind of metadata that takes the information of the image with the help of different tags, keywords, or descriptions of the image. The word content of CBIR refers to color, shape, texture, or some other information that is useful for the image description. Searching on metadata depends on consequent choice and completeness. In other words, users manually search a query image by entering keywords in a very large database which takes a long time. The users are also not able to analyze the right query information about an image. This assessment for retrieving the effective information from an image is not well defined. So the CBIR method is used effectively and efficiently which

faces similar challenges.

The content-based image retrieval system has two basic challenges: first is the intention gap and the second is the semantic gap. In the intention gap, related difficulties of users are the suffering to precisely express the image content by a query, like an image or a sketch map. The semantic gap is related to the origin of the problem which is used to describe the tall height image features with the semantic concept of short height image features [1 - 3].

Content-based image retrieval, introduced in 1992 by T. Kato, explains the experimental method to retrieve image features like color or shape features from the huge image database automatically. The content-based term used here describes the method of obtaining images from a image features database, with various methods, tools, and algorithms that are designed to develop the areas like statistical methods, object detection, pattern recognition, and computer vision, etc.

In the early 1920s, some new research methods in CBIR are introduced. Especially, two types of research are in progress to the present-day for retrieving the

* Corresponding author. Tel.: +90-212-604-0100.

E-mail addresses: h.koyuncu@outlook.com (H. Koyuncu), manishdixit@ieee.org (M. Dixit), baki.koyuncu@altinbas.edu.tr (B. Koyuncu)

ORCID: 0000-0002-8444-1094 (H. Koyuncu), 0000-0003-2589-6010 (M. Dixit), 0000-0002-0507-3431 (B. Koyuncu)

DOI: 10.35860/iarej.811927

This article is licensed under the CC BY-NC 4.0 International License (<https://creativecommons.org/licenses/by-nc/4.0/>).

images from the large multimedia databases. In the primary research, the local image feature of Scale Invariant Feature Transform (SIFT) is contented [4]. SIFT is an image feature detection algorithm that detects and describes local features. The "Bag of Visual Words" (BOW) model [5] is used in secondary studies. The BOW model makes a powerful summary of demonstration or illustration of images based on the quantization of the restricted local features and promptly modifies the characteristic of file indexing of an organization for image retrieval. A recent Literature survey on multimedia content-based image retrieval is presented as a reference in this study [6 - 29]. The general contribution of this study was to explain and analyze the collection of CBIR works with their details for the readers.

The manuscript is organized as follows; section 2 is a literature review, sections 3 and 4 explain Image Retrieval Techniques and Content-Based Image Retrieval Techniques, section 5 presents Content-Based Image Retrieval Methods. Section 6 and 7 explain Distance Measurements and Performance Evaluation. Finally, sections 8 and 9 present "Results and Discussions" and Conclusions.

2. Literature Review

The research by Yue et al. [30] focuses on color and texture low-level characteristics extracted from CBIR. These two types of features are based on a co-occurrence matrix to extract image features to form the feature vectors after being used in a global color histogram, local color histogram, and texture features which are examined by CBIR. After feature extraction, it calculates the Euclidean distance measurements to find the images. CBIR methods are designed by applying color and texture fused features with the help of weight constructions of feature vectors. The process of retrieving the experiments by showing the combination of features is retrieved by bringing an enhanced illustration of the single feature. This proceeds to a superior retrieve outcome. The whole work is done on Java Eclipse enlargement atmosphere and SQL server 2005 is used as the database system.

Singh et al. [31] in their study, solve the content-based image retrieval in energetic surroundings with address problems. It focuses on implementing a new structure that is capable of search and correct features to examine the new query images that improve retrieval accuracy and to make it more effective. It works on the Fuzzy C-Means (FCM) algorithm that generalizes the hard C algorithm. It is produced as a soft panel in a set database. The invariant array vector of images is extracted using a feature extraction tool such as Fast Fourier Transform in this research. Aside from the HSI component of the color

image, the resulting array vector is used as a first feature vector after image segmentation, and then a second feature vector is used. The study proposed an algorithm on 100 tested different images and produce a better performance as compared to the traditional method of CBIR.

Alsmadi[32] developed in his study, a novel relationship between the evaluation of a heuristic algorithm known as a memetic algorithm (Genetic algorithm with a great deluge) to achieve features of the target image and the train image. It comprises of Gradient Descent (GD) algorithm with a Genetic algorithm (GA). It also increases the quality of weights (solution) by increasing the fitness function (number), which helps in the process of searching. Hence, it tests comparable images that are retrieved from a large image database and the calculated results work as a measure of average precision and recall rate performance. Filtering processes can be used in the CBIR in the future to provide more precise results.

Ahamed et al. [33] focus on accessible CBIR system limitations like bandwidth requirement, data security, and storage space, to defeat these limitations. A new technique, CBIS prediction Errors, (CBIR-PE), is presented that uses to rectify errors instead of actual images for storage, transmission, and retrieval. It proposes new techniques that are based on groupings like the clustering technique called WBCT-FCM, WBCT, and FCM. The performance of the proposed WBCT-FCM and CBIR-PE are evaluated using the COREL-1k database. After testing, results are much better as compare to the traditional clustering technique. It retrieves better accuracy as compared to the existing methods.

Saeed et al. [34] proposed a new technique of edifice feature vector to represent images for clustering. It consists of 140 elements which take different features such as color histogram, color moments, Gabor filters, GLCM matrix, etc. It works implicitly on the core database which contains 1000 colored images.

In the study of Joshia et al. [35], the results are deployed from binary and grayscale image retrieval. It is detected by descriptors for color IRS, while it does not detect perfect query. It uses the KCOLavg descriptor, which eliminates the combination given by three colors and considers the average color contribution in the background. Therefore, it does not give a better efficient descriptor in color images. Hence, the study of the intuitive descriptor gives a better output.

Sharma et al. [36] focus on feature extraction and homogeneous attribute measures, utilizing a pyramid that is prepared on wavelet decay and energy level calculations. These energy levels are equivalent to the manipulative distance between the target image and the training images. A substantially huge image database

from the Brodatz album is applied for retrieval claims. The investigational report shows that the prep underrating of Canberra, Bray-Curtis, Square chord, and Square Chi-squared distances are more than the conservative Euclidean and Manhattan distance.

Mistry et al. [37] proposed a hybrid feature system for competent CBIR. It is based on spatial, frequency, Color, and Edge directivity Descriptor (CEDD), and BSIF feature descriptors. The proposed method uses the WANG database; it contains 1000 Corel images of 10 various types of .jpg images, of size 384×256 or 256×384 . It combines 100 images in 10 various groups like animals, vehicles, objects, places, etc.

Khodaskaret al. [38] presented a study on CBIR use in the color feature. The likelihood of histograms for each color component is increased with this feature. These histograms are separated into different numbers of major coordinates and for each coordinate, different statistical features like standard deviation, skewness, mean, variance, kurtosis, etc are calculated. The processing rate is slow, and the proposed technique is deployed on average image databases achieving better results.

Pradeep et al. [39] demonstrated the CBIR approach by combining artificial neural networks, fuzzy logic, support vector machines, and other soft computing methods. The traditional CBIR system retrieves images by using low-level image features, and it gets the semantic gap. It suggests an advanced structure for CBIR that increases the accuracy of image retrieval by combining fundamental soft computing methods. The proposed system works with relevant feedback based on a Support vector machine (SVM). It gives an intelligent classification of images based on a relevant or irrelevant query image, and later, it calculates the performance measures like precision, recall, and accuracy.

Katira et al. [40] presented the technique of CBIR, which is used with the ordered-dither block truncation coding, (ODBTC), for the image content descriptors. The encoding is done by using ODBTC that combines the image blocks into quantized and bitmap images, such as color co-occurrence features, (CCF) and bit pattern features, (BPF). This is converted into ODBTC encoded data streams without any decoding. The experimental result shows that the planned scheme is greater than the block processing coded image retrieving systems.

Gupta et al. [41] proposed some primitive features of an image that were utilized in the current scenario system. Some features are selected based on similarity identification between the images. The proposed algorithms are used for calculating the similarity between image features and then describe them. MATLAB tool is used for image verification in a database.

Wavelet-Based Color Histogram Image Retrieval was proposed by Bagri et al. [42] for feature selection such as

color and texture (WBCHIR). In the path of wavelet transformation and color histogram, shape and shade features are often used for feature selection, making these features energetic and adaptable. In this study, the segmentation and grid information is used for the first time for feature extraction.

Giveki et al. [43] used the texture Gray Level Co-occurrence Matrix, Hue moments, and the grouping of Tamura texture features and shape invariant Hue moments to compare the combination of texture and shape features. The system's efficiency is measured using the accuracy and recall methods.

Selviet al. [44] presented experimentation on the effectiveness of choosing color space performance using CBIR for Wavelet decomposition. Hence, the efficiency increased in the results of CBIR representation using wavelet transform in color space and color moments.

Sasikala et al. [45] proposed three algorithms to increase the performance. In this study, the Heterogeneous Minimum Order k-SimRank (HMok-SimRank) algorithm is used which derives the corresponding algorithm using Integrated Weighted Similarity Learning (IWSL) to integrate the data for query images in an image database. A ranking algorithm is used for finding the rank of the images which increases both relevance and speed.

Singh et al. [46] introduce CBIR problems and challenges, by using an accurate image searching system. It is based on CBIR using Mpeg-7 descriptors, which works on low-level features and is extracted from the large database. it transforms between color spaces and quantization and uses color information for feature extraction. This proposed method achieves robustness and 92.4% accuracy.

Dharani et al. [47] proposed a survey study on the CBIR technique. The study searches the gigantic image database for the target image based on the user query using visual characteristics of an image such as color, shape, texture, and so on. It is based on labeled and unlabelled images to analyze the efficiency of an image using the method like D-EM, SVM, RF, etc.

Smeulders et al. [48] discuss the current state of content-based retrieval from various image formats, as well as the role of semantics, pattern, and the sensory gap. A special retrieval strategy, such as global features, object and shape, salient points, structural combinations, and signs, is used for image and object similarity searches. This is accomplished through the use of system engineering databases, system design, and evaluations.

Li et al. [49] focus on the number of digital images increasing from time to time and taking out from huge databases which is very complex. Indexing image data is based on the text. The indexing uses low-level features of the image that may diminish the workload and make

mining to become more rapid. The indexing technique index the digital images in the database by the peak color proportion. The images will be routinely confidential by their low-level feature like color. Execution of this technique will benefit image mining.

Tarulatha et al. [50] work on advanced technologies to increase accuracy using CBIR systems. The study centers on the shifted design which is complicated for the low-level feature extract algorithms that reduce the semantic gaps among image features and databases. The authors carried out a literature survey on recent technologies to achieve an elevated level of semantic-based image retrieval.

Liu et al. [51] focused on Gabor wavelet proof texture analysis for image retrieval method based on the Gabor filter. The texture feature works on the mean and variance of statistical features for Gabor filtered image. In normalization, it is used by a spherical shift of the feature elements which override on points. The image indexing and retrieval are conducted on the image texture and standard images.

Siradjuddin et al. [52], in Content-based Image Retrieval, provided an autoencoder using a Convolutional Neural Network for feature extraction. In the convolutional autoencoder architecture, the encoder and decoder layers are used. The encoder layer reduces the image's dimension by applying the convolutional neural network's feature learning capability to it. The decoding layer rebuilds the autoencoder's output as closely as possible to the data input. The results found that the extracted features can be used to represent images and recover relevant images in content-based image retrieval.

Fadaei et al. [53] suggested a technique that extracts the Zernike moments from the query image and calculates an interval for that query. The retrieval process ignores images in the database that are beyond the interval. As a result, a database reduction happens before retrieval, resulting in increased speed. Relevant images for query images are stored in the reduced database, while irrelevant images are discarded.

Naoufal et al. [54] used a procedure based on color features coded as strings and genetic algorithms to create a content-based image retrieval system. The goal is to find the fittest people in a huge group of people. After that, they compare the results achieved by varying the size of the images at different intervals to see how the size affects their technique to reduce the computation time. This method claims to lower the cost of retrieving other features while also improving retrieval accuracy.

Rudrappa et al. [55] showed a system that uses ground-based imagery of clouds in the sky to determine if they are low, middle, or high-level clouds. For cloud classification, they use K-means clustering and Content-Based Image Retrieval (CBIR) approaches. The evolved

system divides clouds into three categories: low, middle, and high-level clouds. The outcome of this cloud identification can then be used as an input to a system that determines rainfall dynamically.

Finally, in the commercial field, The CBIR method was implemented by IBM, known as QBIC (Query-based Image Content). QBIC works on networking and graphics-based approaches. It is a simple, well-understood, and attractive substitute to traditional methods. A CBIR system architecture is presented in Figure 1.

The offline stage and the online stage are the two steps of a common content-based image retrieval framework. In the offline stage, the image is captured and placed in the image database. Later on, images are represented, and data indexed. In the online stage, other query methods, image features, image feature advantages, and disadvantages of features are considered for image retrieval. CBIR Framework module is given in Figure 2.

3. Image Retrieval Techniques

The image retrieval system is a structure that allows you to search, discover, and retrieve images from a large image database. Figure 3 shows a block diagram of the image recovery system. Traditionally, image retrieval uses some common methods for adding metadata like capturing, keyword, or some other type of information that describes an image. Early research works on image retrieval techniques are summarized here for the reader's attention.

3.1 Text-Based Image Retrieval

The text-based image retrieval method also known as the descriptive image-based system is used for retrieving XML documents that contain images.

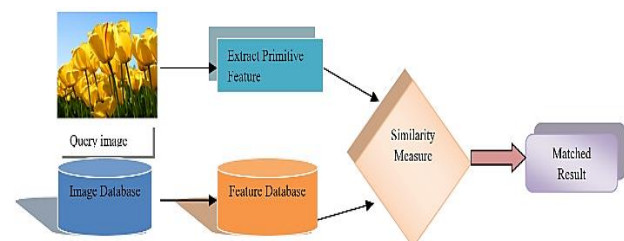


Figure 1. CBIR system architecture

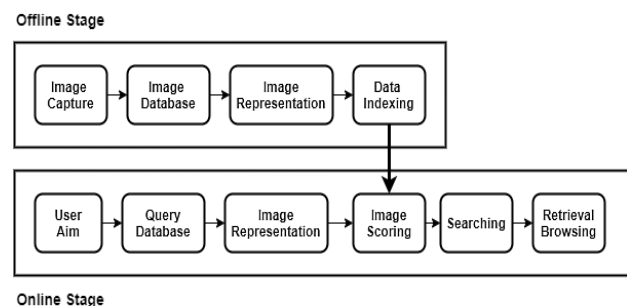


Figure 2. CBIR framework module (online/offline stages)

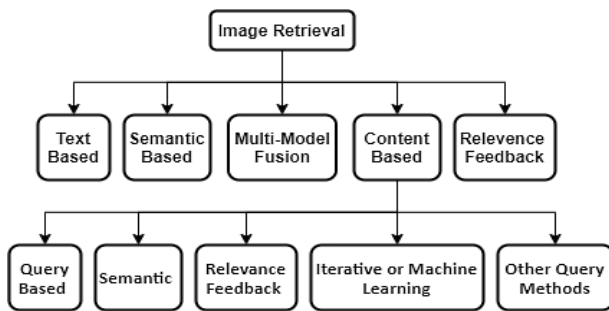


Figure 3. Block diagram of Image retrieval technique

It is based on textual information for an exact multimedia query in the database. Therefore, it overcomes the limitation of the CBIR system. It allows the users to present information according to the needs of textual query and finding the significant images which match with the textual query and the database [56-60].

3.2 Multimodal Fusion Image Retrieval

Multimodal fusion image retrieval is required for data synthesis and a machine learning algorithm, called, combination of evidence is deployed. This algorithm is used for merging multiple sources of evidence, which is based on multiple modalities, like the skimming effect, chorus effect, and dark horse effect [61].

3.3 Semantic-Based Image Retrieval

A semantic-based image retrieval system is used for current images. It is the method, that works to solve the semantic gap problem, which leads to two main approaches. The first is Annotating images and the second is image segmentation that uses automatic image explanation and adopting the semantic web initiatives [62, 63].

3.4 Relevance Feedback Image Retrieval

It is used to determine the distinction between user information and image representation, which determines the limit of nuclear retrieval systems' image retrieval accuracy. It's used to fill in important semantic gaps. The underlying concept behind relevance feedback is to incorporate the user's perception of the query and have them compute the retrieval results. For better results, user interaction is required [64 - 67].

4. Content-Based Image Retrieval Techniques

CBIR systems are projected on many techniques. But there are some problems with the image retrieving system based on pixel contents [68-72]. Early research works on content-based image retrieval techniques are summarized here for the reader's attention.

4.1 Query-Based Technique

Query by example is based on a query technique,

which works based on a searching algorithm. The searching algorithm is based on the functionality of the resulting image and shares common image information which is known as Reverse searching image. Some common examples of an image searching system are as follows:

- An existing image can supply by the users selected from a random set.
- The user draws a rough calculation of an image.

4.2 Semantic Retrieval

Semantic retrieval is used to start a user quest, for example, "find pictures of Abraham Lincoln". It is a sort of open-ended task, which is extremely complicated for computer operations. Lincoln cannot face the cameras and can't use the same poses. CBIR systems use low-level features such as texture, color, and shape features. These features are used in grouping and interfacing so that it is easy to input the databases which are qualified to match the features such as the face, fingerprint, or shape. Generally, image retrieval requires a human identification of the higher-level concept.

4.3 Relevance Feedback (Human Interaction)

The Combine CBIR system is a searching technique for the identification of the large variety of probable user images and the target can be a strong task. A CBIR system relies totally on the understanding ability of user queries.

4.4 Iterative Or Machine Learning

Machine learning and application is an iterative technique for the fitting of extra culpability in CBIR.

4.5 Other Query Methods

These methods are based on searching and browsing, for example, navigation of modified and hierarchical categories such as query by image type, query by a various instance of images, query by visual outline, query by the straight requirement of image features, and multimodal queries (e.g. combining touch, voice, etc.)

5. Content-Based Image Retrieval Methods

Content-based image retrieval (CBIR), which is used as a distance measure, is the most common technique for comparing two images. The distance is computed by comparing the two images' color, texture, and shape similarity.

5.1 Color

One of the most important features of a content-based image retrieval system is the color feature, which aids image recognition. Color features are a pixel property and

these properties work on the reflection of light. Color tells about the variation between object, position, and time of day. There are many types of color model that describes the color information of an image, and it reduces image information about potential inequity to the single gray level values.

Color space is a method that represents pixel coordinates in 3D color space, in which the popular Red-Green-Blue (RGB) method is used for image retrieval. RGB method consists of three colors which are red, green, and blue, together with additive primaries, which are developed by combining all three colors. In contrast, the cyan-magenta-yellow (CMY) method is a subtractive color model and is used mostly for printing. It is developed for brightness inclusion.

Both the color models RGB and CMY are mechanism-based or perceptually non-uniform. The CIE (International Color Commission) measured the colors and perceptually standardized them. CIE is made for subtractive color mixtures and designed the color management [47].

The Hue, Saturation, and Value (HSV) space model is widely used in computer graphics. The invariant moment is used to change the light and camera direction, and it is an appropriate method for image retrieval. The invariant moment is represented by calculating the wavelength of light, and it expresses the pure spectrum color model which ranges from 0° to 360°. RGB coordinates are easily translated into HSV coordinates and are represented by using equations (1), (2) and (3). The conversion of the RGB space color model to the HSV space image model is shown in Figure 4. Example color conversions to a different color model are presented in Figure 5a and Figure 5b.

$$H = \cos^{-1} \frac{[(R - G) + (R - B)]}{\sqrt{[(R - G)^2 + (R - B)(G - B)]}} \quad (1)$$

$$S = 1 - \frac{3[\min(R, G, B)]}{R + G + B} \quad (2)$$

$$V = \frac{R + G + B}{3} \quad (3)$$

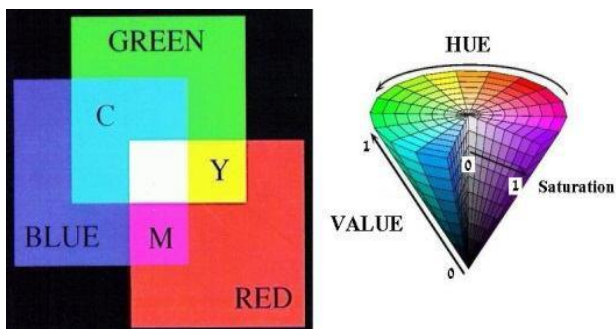
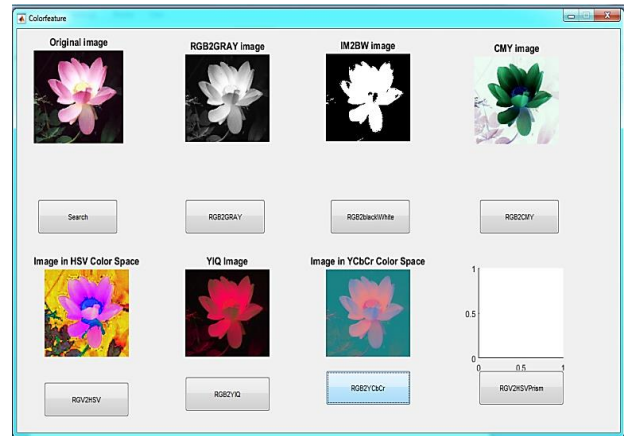
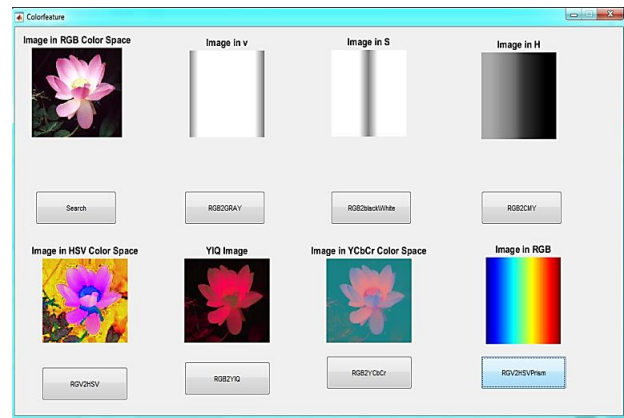


Figure 4. Block diagram of image retrieval technique



(a)



(b)

Figure 5. Color conversion to a different color model

Some common techniques used in the color feature are as follows:

5.1.1 Color moment

The most popular color feature used for CBIR is the color moment. The color moment is a statistical measure that differentiates images using the feature extraction method and that image is recognized by the distribution of color moment (statistical measure) or features. There are some common statistical measures namely mean, variance, and skewness shown by equations (4), (5), and (6).

The color moment proves the efficiency and effectiveness of color statistical distribution of images. A sample of the image database used in calculations is presented in Figure 6.

$$\mu_i = \frac{1}{N} \sum_{j=1}^N f_{ij} \quad (4)$$

$$\sigma_i = \left(\frac{1}{N} \sum_{j=1}^N (f_{ij} - \mu_i)^2 \right)^{\frac{1}{2}} \quad (5)$$

$$s_i = \left(\frac{1}{N} \sum_{j=1}^N (f_{ij} - \mu_i)^3 \right)^{\frac{1}{3}} \quad (6)$$

Table 1. The results of color moment calculations for the database in Figure 6

No.	ColorMo1	ColorMo2	ColorMo3	ColorMo4	ColorMo5	ColorMo6
Image1	98.5513	150.7495	101.2073	85.5941	101.1767	77.3031
Image2	95.9414	64.8982	58.3551	83.2036	57.7118	87.1210
Image3	69.7427	111.2991	73.9050	84.5222	73.8894	48.2561
Image4	77.0045	33.9630	50.8406	67.5750	50.2813	53.6986
Image5	82.2800	125.5592	67.4838	67.3654	67.6743	60.8554
Image6	84.2418	50.5414	43.1666	61.5602	42.4683	70.5261

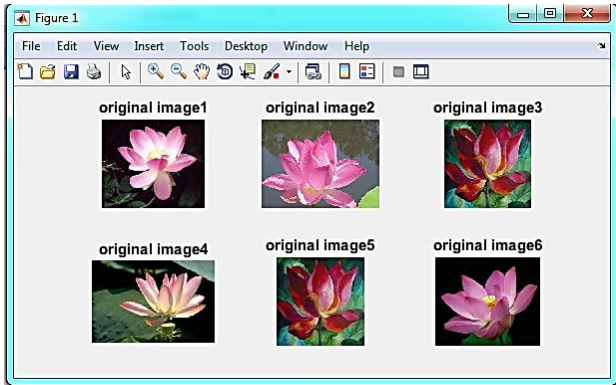


Figure 6. Sample of image database

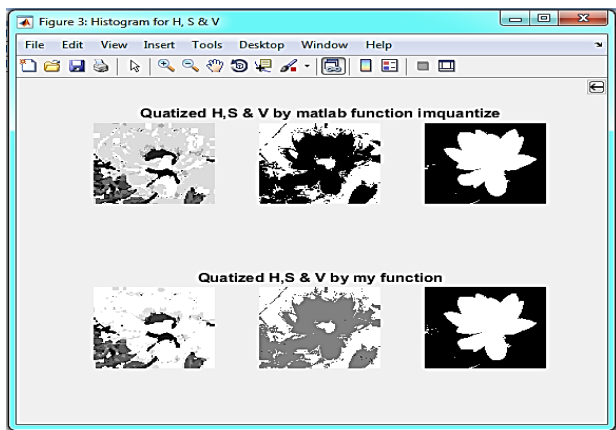


Figure 7. HSV histogram for color model

where f_{ij} is the image value, i^{th} is the color element of an image pixel, j and N are the numbers of the pixel images.

Color moment calculations are carried out for the 6 original images in Figure 6 and the results are presented for the color moments in Table 1. The images were obtained from an online database atozflowers.com/. A custom software was written in Matlab to calculate the color moment presented in Table1.

5.1.2 Color Histogram

A color histogram is an interpretation and identification of RGB image intensity. Color histogram defines a color vector, which is a set of pixels of all the images, coordinates which are denoted by the pixels, and the images considered individually as an RGB image. It is a way of viewing translation and rotation, by changing an image slowly with the help of scalar transformations and viewing angles. H is a color histogram vector for an image defined in equation (7) as:

$$H = \{H[1], H[2], H[3], \dots H[i], \dots H[N]\} \quad (7)$$

where i is a color histogram, $H[i]$ is the number of pixelscolor, i is an image, and N is the number of the coordinate for a color histogram. An example HSV histogram for the color model is presented in Figure 7.

5.1.3 Color Correlogram

The color correlogram is a color-based table indexing feature, with the k^{th} entry for (i, j) indicating the possibility of indexing color j coordinates at a distance of k from a pixel of color I in the image.

The mathematical representation of color correlogram is shown in equation (8).

$$Y_{(i,j)}^{(k)} = Pr_{p1 \in C(i), p2 \in I} [P2 \in I_c(j) | P1 - P2 | = K] \quad (8)$$

where $i, j \in \{1, 2, \dots, N\}$, $k \in \{1, 2, \dots, d\}$, and $|p1 - p2|$ is the distance between pixels $p1$ and $p2$.

5.1.4 Color Coherence Vectors

The color coherence vector is divided into two categories namely coherent histogram and non-coherent pixel. Pixels are measured to be coherent if they are part of a continuous equally colored area and the size of this area exceeds some threshold. This area is usually defined as 1% of the image area. Finally, the general advantages and disadvantages of color models are summarized as shown in Table 2.

Table 2. Advantages and disadvantage of color models

Models	Advantages	Disadvantages
RGB (Red, Green, Blue)	The transformation method is not used for displaying information, as it is measured as the basis of color space for the different methods, it uses video display because of its stabilizer properties.	It is not used in object identification and recognition of colors.
CMY (K)	It is normally used for the production of printer color.	It is a subtractive model so inks are not given as color output.
HSV	HSV color model can be simply identified by individual observation as compare to CMY or RGB.	Hue coordinates are receptive to derivations because of their sharp nature features.

5.2 Texture

The texture of an image is its most important feature. It is a set of the matrix which calculates the spatial collection of color intensities of an image and selects the region of an image. It can be synthetically created, and it identifies natural scenes that are captured in an image. Segmentation, classification, pattern recognition, and machine vision all use the texture feature. Some of the most prevalent approaches are directionality, contrast, coarseness, roughness, regularity, and line-likeness. These are divided into three approaches: structural, statistical, and spectral.

Structural approach: It is an image texture, a set of primitive texture elements (pixel or texels), in morphological operators and adjacency graph. It is based on a regular subpattern.

Statistical approach: It is a quantitative measure, like co-occurrence matrices, Tamura feature, Fourier power spectra, and shift-invariant principal component analysis

(SPCA), etc.

Spectral approach: It is a property used for the Fourier spectrum, which explains the global periodicity of a gray level image by the identification of high energy coordinates in the Fourier spectrum.

The commonly used techniques in the statistical approach are:

5.2.1 Gray Level Co-Occurrence Matrix

It is developed by R.M. Haralick, and it is the most important and popular method for representing image textures. It is the second-order statistical method in which 14 statistical measures are used for feature extraction. See Table 3. GLCM Texture Analysis examples with different pixel distances are given in Figures 8,9 and 10.

GLCM Texture feature analysis is carried out with the images in the image database in Figure 6. The results are presented in Table 4.

Table 3. Advantages and disadvantage of color models

No.	Feature	Formula	Remark
1	Contrast	$\sum_{n=0}^{Nn-1} n^2 \{ \sum_{i=1}^{Ng} \sum_{j=1}^{Ng} P(i, j) \}, i - j = n$	Have a discriminating ability. Rotationally variant
2	Entropy	$-\sum_i i \sum_j P(i, j) \log(P(i, j))$	Have a strong discriminating ability. Almost rotational invariant.
3	Correlation	$\frac{\sum_i i \sum_j j P(i, j) - \mu_x \mu_y}{\sigma_x \sigma_y}$	Have a strong discriminating ability. Rotational dependent feature.
4	Sum average	$\sum_{i=2}^{2N} i p_{z+y}(i)$	Characteristics are similar to variance. Rotation is invariant.
5	Sum of squares: Variance	$\sum_i i \sum_j j (i - \mu)^2 P(i, j)$	Computationally expensive Rotation variant.
6	Info. The measure of Correlation 1	$\frac{HXY - HXY}{\max \{HX, HY\}}$	Similar to the Correlation
7	Info. The measure of Correlation 2	$(1 - \exp[-2(HXY2 - HXY)])^{\frac{1}{2}}$	Similar to the Correlation
8	Inverse Different Moment	$\sum_i i \sum_j j \frac{1}{1 + (i - j)^2} P(i, j)$	Similar to the angular second moment.
9	Sum Variance	$\sum_{i=2}^{2N} (i - f_s)^2 P_{(x+y)}(i)$	Similar to variance
10	Sum Entropy	$-\sum_{i=2}^{2N} P_{(x+y)}(i) \log\{P_{(x+y)}(i)\} = f_s$	Similar to entropy
11	Angular second moment / Energy	$\sum_i i \sum_j j P(i, j)^2$	No distinguishing ability
12	Difference variance	$\sum_{i=0}^{N-1} i^2 P_{(x-y)}(i)$	Similar to Variance
13	Difference Entropy	$-\sum_{i=0}^{N-1} P_{(x-y)} \log\{P_{(x-y)}(i)\}$	Similar to Entropy
14	Max. Correlation coeff./ Mean	$Q(i, j) = \sum_k k \frac{P(i, j)P(j, k)}{p_x(i)p_y(k)}$	The square root of the second-largest Eigenvalue

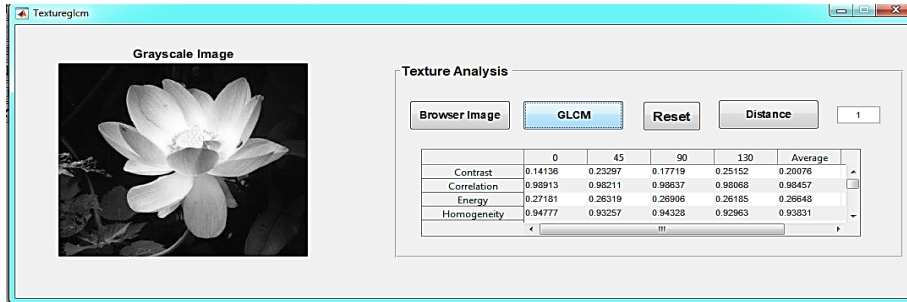


Figure 8. Texture Analysis for 1-pixel distance

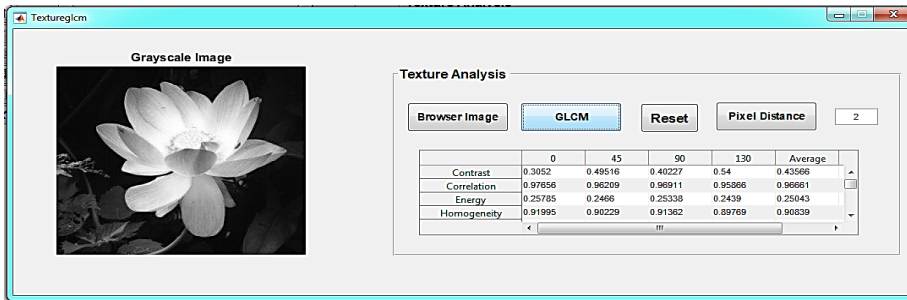


Figure 9. Texture Analysis for 2-pixel distance

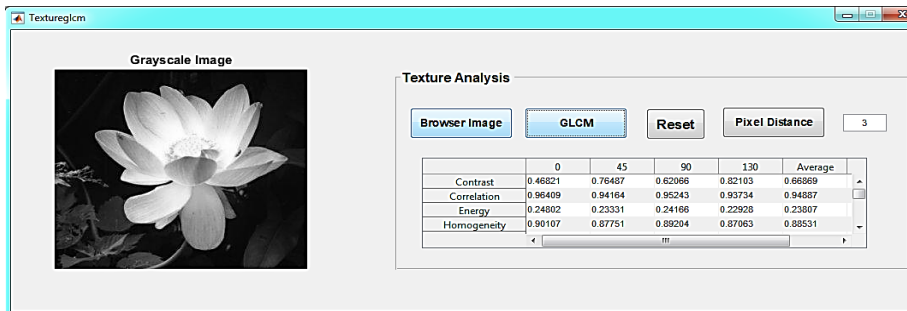


Figure 10. Texture Analysis for 3-pixel distance

5.2.2 Tamura texture

Tamura feature is designed by studying the human perceptual experience of texture features. There are some perceived texture feature types which are as follows:

- Coarseness
- Contrast
- Directionality
- Linelikeness
- Regularity

Coarseness:

It is a distance measure related to the spatial variation of a gray-level image. Hence it is easy to find the size of the primitive element (pixels) from the texture feature. Coarseness is the quantitative measure of a scalar image with a fixed window size. The smaller value of the texture element is known as coarser.

Contrast:

It is a measure to check the image whether is black or white based on distribution.

Directionality:

It is the information of edge strength and directional angle. It uses the pixel-wise differential for computing an image.

Linelikeness:

It describes the average conjunction of edge directions, so it defines the isolated pixels by using distance and with the direction α .

Regularity:

It defines coarseness, contrast, and direction. It is denoted by users as a normalization factor and σ defines the standard deviation of the texture features of each sub-image.

5.2.3 Wavelet transform

Wavelet transforms offers the multiresolution method that describes texture analysis and classification. It explains the basic function of a superposition in the family called wavelets. It uses a multi-resolution approach to calculate the 2D image. It employs recursive filtering and sub-sampling, which divides the image into sub-levels. Every level is decomposed into sub-bands of four frequencies which are, LH, LL, HH, and HL. L represents low frequency and H denotes the high frequency. Haar, Coiflet, Mexican Hat, Morlet, and Daubechies are some examples of wavelets. The most common and basic wavelets is Haar wavelet, and Daubechies is a fractal structure application.

Table 4. Texture features of GLCM

No.	Feature	Image1	Image2	Image3	Image4	Image5	Image6
1	Autocorrelation 1	15.6346	15.6346	15.6346	15.6346	15.6346	15.6346
2	Autocorrelation 2	15.6675	15.6675	15.6675	15.6675	15.6675	15.6675
3	correlation 1	0.9691	0.9691	0.9691	0.9691	0.9691	0.9691
4	correlation 2	0.9765	0.9765	0.9765	0.9765	0.9765	0.9765
5	Cluster Prominence 1	1303.913	1303.913	1303.913	1303.913	1303.913	1303.913
6	Cluster Prominence 2	1311.307	1311.307	1311.307	1311.307	1311.307	1311.307
7	Cluster Shade 1	100.7474	100.7474	100.7474	100.7474	100.7474	100.7474
8	Cluster Shade 2	101.3129	101.3129	101.3129	101.3129	101.3129	101.3129
9	Dissimilarity 1	0.2155	0.2155	0.2155	0.2155	0.2155	0.2155
10	Dissimilarity 2	0.1883	0.1883	0.1883	0.1883	0.1883	0.1883
11	Max Probability 1	0.4785	0.4785	0.4785	0.4785	0.4785	0.4785
12	Max Probability 2	0.4827	0.4827	0.4827	0.4827	0.4827	0.4827
13	Sum of Square 1	15.7321	15.7321	15.7321	15.7321	15.7321	15.7321
14	Sum of Square 2	15.7321	15.7321	15.7321	15.7321	15.7321	15.7321
15	Sum of Average 1	6.10709	6.10709	6.10709	6.10709	6.10709	6.10709
16	Sum of Average 2	6.10288	6.10288	6.10288	6.10288	6.10288	6.10288
17	Sum of Variance 1	42.8879	42.8879	42.8879	42.8879	42.8879	42.8879
18	Sum of Variance 2	43.0304	43.0304	43.0304	43.0304	43.0304	43.0304
19	Sum of Entropy 1	1.9545	1.9545	1.9545	1.9545	1.9545	1.9545
20	Sum of Entropy 2	1.9435	1.9435	1.9435	1.9435	1.9435	1.9435
21	Diff of Variance 1	0.4022	0.4022	0.4022	0.4022	0.4022	0.4022
22	Diff of Variance 2	0.3051	0.3051	0.3051	0.3051	0.3051	0.3051
23	Diff of Entropy 1	0.5566	0.5566	0.5566	0.5566	0.5566	0.5566
24	Diff of Entropy 2	0.5135	0.5135	0.5135	0.5135	0.5135	0.5135

5.3 Shape

The shape, which is used to extract information from images, is the most important feature. Because of noise, occlusion, and arbitrary distortion, the object recognition problem becomes more complicated. Shape representation is focused on the shape features and it is a boundary-based or region-based shape. To be more effective, a shape feature must have the following properties: identifiability, translation, occultation invariance, noise resistance, affine invariance, reliability, and statistical independence.

5.3.1 Boundary-based shape

This feature mainly focuses on the outer boundaries of an image that defines the important region by using the external characteristic of an image, in each pixel (or coordinate) along the image boundary. See Figure 11. Some representations of Boundary-based shapes are as follows:

1. Polygonal Models, or Boundary partitioning
2. Fourier Descriptors
3. Spines, higher-order constructions
4. Curvature Models

5.3.2 Region-based shape

This feature focuses on the shape region of the whole

image by using a complete image internal description. some representations of Region-based shapes are as follows:

1. Super quadrics
2. Fourier Descriptors
3. Implicit Polynomials
4. Blum's skeletons

5.3.3 Fourier descriptor

Fourier descriptor describes the shape feature of an object in an image, in which it uses Fourier transform for boundaries and considers the shape (contour) of a two-dimensional image. Contour representations are divided into three types:

Curvature: Curvature $K(s)$ is a method that changes the rate of the tangent direction of the curve, it is represented as s , s is the point of the curve. It can be represented as:

$$K(s) = \frac{d}{ds} \theta(s) \quad (9)$$

where $\theta(s)$ represent the turning function of the curve.

Centroid distance: Centroid distance is a method that calculates the distance function between the boundary pixel and the centroid (x_c, y_c) of an image object, It's defined as:

$$R(s) = \sqrt{(x_s - x_c)^2 + (y_s - y_c)^2} \quad (10)$$

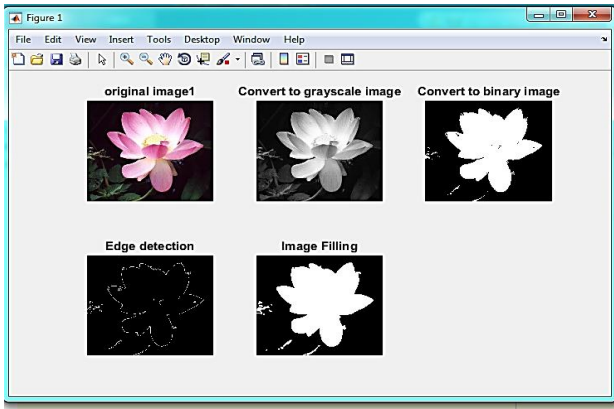


Figure 11. Edge detection

Complex Coordinate: Complex coordinates describe the coordinates of boundary pixels and the complex number is defined as:

$$Z(s) = (x_s - x_c) + j(y_s - y_c) \tag{11}$$

5.3.4 Fourier transforms

Contour representation is used to generate the complex coefficient representation of the object shape in the frequency domain. It can be divided into three groups. The first one is the low frequency coefficients that express the shape property. The second one is the higher frequency coefficients that reproduce the shape information. The third one is the complex frequency coefficients which achieve the rotation invariant, scalar invariance, amplitude coefficients separated by the amplitude of the DC component, and non-zero coefficients. The transformation invariance is based on the contour illustration.

5.3.5 Moment invariant

It is based on conventional shape representation. It is a set of moment invariants. The main reason to propose the moment invariant is to calculate the region-based moment. if the object R is represented as the binary image, then the center moment order of p+q for the shape of object R is defined as:

$$\mu_{p,q} = \sum_{(x,y) \in R} (x - x_c)^p (y - y_c)^q \tag{12}$$

where (x_c, y_c) represents the center of the image.

Invariants are independent of position, size, and orientation. Moment invariant is the basis of these moments. These are translation; rotation and scale and can be derived as:

$$\Phi_1 = \mu_{2,0} + \mu_{0,2} \tag{13}$$

$$\Phi_2 = (\mu_{2,0} - \mu_{0,2})^2 + 4(\mu^2)_{1,1} \tag{14}$$

$$\Phi_3 = (\mu_{3,0} - \mu_{1,2})^2 + (3(\mu_{2,1}) - \mu_{0,3})^2 \tag{15}$$

$$\Phi_4 = (\mu_{3,0} + \mu_{1,2})^2 + (\mu_{2,1} + \mu_{0,3})^2 \tag{16}$$

$$\begin{aligned} \Phi_5 = & (\mu_{3,0} - 3\mu_{1,2})(\mu_{3,0+\mu_{1,2}})[(\mu_{3,0} + \mu_{1,2})^2 \\ & - 3(\mu_{2,1} + \mu_{0,3})^2] + (3\mu_{2,1} \\ & - \mu_{0,3})(\mu_{2,1} \\ & + \mu_{0,3})[3(\mu_{3,0} + \mu_{1,2})^2 \\ & - (\mu_{2,1+\mu_{0,3}})^2] \end{aligned} \tag{17}$$

$$\begin{aligned} \Phi_6 = & (\mu_{2,0} - \mu_{1,2})[(\mu_{3,0} + \mu_{1,2})^2 - (\mu_{2,1} + \mu_{0,3})^2 \\ & + 4\mu_{1,1}(\mu_{3,0} + \mu_{1,2})(\mu_{2,1} \\ & + \mu_{0,3})] \end{aligned} \tag{18}$$

Hue moments for the 6 test images in Figure 6 are tabulated in Table 5.

5.3.6 Turing angles

In contour representations of 2D images, the close sequence representation leads to successive boundaries of the pixel (x_s, y_s) , where $0 \leq s \leq N-1$ and N define the total number of the element on the boundaries. The turning function is represented as $\theta(s)$, which finds the angle of the counter. It uses a clockwise tangent for the functions of the curve length s according to the situation point on the image (curve). It is defined as

$$\theta(s) = \tan^{-1} \left(\frac{y'_s}{x'_s} \right) \tag{19}$$

$$y'_s = \frac{dy_s}{ds} \tag{20}$$

$$x'_s = \frac{dx_s}{ds} \tag{21}$$

A summary of various Feature techniques is presented in Table 6.

Table 5. Hue moments for 6 test images in Figure 6

No.	Hue Moment	Image1	Image2	Image3	Image4	Image5	Image6
1	Moment 1	0.0037	0.0028	0.0034	0.0045	0.00343	0.0051
2	Moment 2	1.1320	6.5133	7.9582	1.8011	7.9544	2.2833
3	Moment 3	6.6040	2.0469	1.6069	2.3019	1.5979	2.8695
4	Moment 4	6.0669	2.0748	1.7688	2.3844	1.7525	2.8935
5	Moment 5	3.8165	4.2588	2.9822	5.5861	2.9328	8.3352
6	Moment 6	2.0412	5.2940	4.9287	1.0061	4.8795	1.3779
7	Moment 7	-4.2586	3.8194	1.6999	-2.2187	1.5708	-2.0836

Table 6. A summary of various feature techniques

Types of feature	Methods	Characteristics
Color Features	Conventional color Histogram(CCH), Fuzzy Color histogram (FCH), [20,21] Color Correlogram	Simple to calculate. It cannot instruct the spatial info. It cannot instruct by color coordinates seminaries. It considers the degree of color relationship between the coordinates. It is strong to quantize fault. Here, the robustness changes the brightness intensity.
Texture Features	Steerable pyramid, Contourlet transform, Complex directional filter bank (CDFB)	In this, we use basic filters for translation and rotation. These filters are the linear combination of basic functions. Only the rotation invariants are the use of texture retrieval. Combining the Laplacian pyramid function with a directional filter requires less computation and difficulty. Optimally, it achieves the joint resolution of space and spatial frequency which is computationally intensive. So, it gives the highest result of texture retrieval. It compares the retrieval outcome with the Gabor wavelet output and it is Shift Invariant.
Shape Features	Fourier Descriptor, Moment Invariants, Directional Histograms	The shape feature is a phase of understanding and implementing the shapes.

6. Distance Measurements

CBIR searches for similarities between the target image and the training images. As a result, the image retrieval method displays the results based on the image list. Many distance measurements are developed for the CBIR feature in recent years. Distance measurements are carried out to locate the connections of feature vectors. In the CBIR system, it is used to compare the relationship between the images. They are briefly summarized as follows:

6.1 Euclidean Distance

It's the measurement of the distance between the image's two coordinates. Euclidean distance is identified as the line segment between 2D Euclidean coordinates X (x₁, x₂) and Y (y₁, y₂). The distance between two coordinates X and Y is defined as shown in equation (22);

$$D(p, q) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \tag{22}$$

where, X = (x₁, x₂, x₃, , x_n) and Y = (y₁, y₂, y₃, , y_n), n is the number of points.

6.2 Standard Euclidean Distance

Standardized data is used to determine standard Euclidean distance. This data can be defined as Standardized Data= (original mean value)/Standard deviation shown in equation (23);

$$d = \sqrt{\sum \left(\frac{1}{s_i^2}\right)(x_i - y_i)^2} \tag{23}$$

6.3 Manhattan Distance

It uses a lattice that is purely based on a diagonally opposite horizontal or vertical path to determine the distance between two points. The Manhattan distance is

computed using Pythagoras Theorem, while the diagonal distance is calculated using the sum of the horizontal and vertical components. Mathematically, it is defined as

$$Mh(p, q) = (x_1 - x_2) + (y_1 - y_2) \tag{24}$$

6.4 Minkowski Distance

It's a popular image retrieval metric in which each image element's feature vector is independent of the others and all features are valued equally. Minkowski distances are used to calculate the distance between two images. Mathematically, it can be defined as

$$D(p, q) = \left(\sum_i (f_i(x) - f_i(y))^p\right)^{1/p} \tag{25}$$

6.5 Mahalanobis Distance

The distance metric is applied to each dimension of the image feature, which is a vector that is unrelated to the others and is applicable in this distance measurement. It can be defined as

$$D(p, q) = \sqrt{(F_p - F_q)^T C^{-1} (F_p - F_q)} \tag{26}$$

where C is representing the covariance matrix of the image feature vector.

6.6 Chebyshev Distance

It calculates the distance between the two points p and q. Chebyshev distance represents the standard coordinates p_i and q_i and it is a metric of the uniform norm. It can be represented as

$$D_{chess}(p, q) = \max(p_i - q_i) \tag{27}$$

where p and q are two points, Cartesian coordinates are (x₁, y₁) and (x₂, y₂).

$$D(i, j) = \sqrt{(F_i - F_j)^T A (F_i - F_j)} \tag{28}$$

where $A(F_i-F_j)$ defines a similar matrix where F_i and F_j are vector lists.

7. Performance Evaluation

Performance measures are used to calculate the results and use image retrieval measures such as precision and recall, which work on feature extraction and give positive stability outcomes. These features extract information without any interpretation problems. It can be categorized as Precision, Recall, and F score.

7.1 Precision

Precision is the performance measure that retrieves the related images to the query image from the total retrieved database images.

$$\text{Precision} = \frac{\text{No. of the relevant retrieved image}}{\text{Total No. of retrieved image}} \quad (29)$$

7.2 Recall

The retrieval of related images to the image database is measured by the recall performance measure.

$$\text{Recall} = \frac{\text{No. of relevant retrieved}}{\text{Total No. relevant in the database}} \quad (30)$$

7.3 F-Score

The accuracy of an image retrieval of the corresponding query image is computed from the database image with precision and recall measurements, and the calculated values indicate the success of image retrieval. These two measurements aren't always enough to maximize the capable image retrieval's correct accuracy. As a result, they were combined into a single value that represents image retrieval accuracy. The F-Score of F-measure is the name for this arrangement. The score is computed using a combination of precision and recall measurements. The weighted average, or harmonic, is what it's called.

$$F = 2 * \frac{(\text{Precision} * \text{Recall})}{(\text{Precision} + \text{Recall})} \quad (31)$$

Finally, comparisons of Feature extraction methods are summarized in Table 7 for the reader's attention.

8. Results and Discussions

Various systematic approaches and general structure of the CBIR scheme are discussed in this review. Techniques for image retrieval are extensively explored and contrasted with the retrieval of content-based images. Content-based methods of retrieval including color structure and form are provided. The important points obtained from the detailed survey carried out in this study are summarized, in Appendix (Table A.1). The survey shows that CBIR is still a very interesting area for research. Different work is undergoing in this field to develop or modernize the CBIR system. It lessens the gaps between the high and low level features. Most of the studies are deployed to optimize retrieval performance. Some recent work on the CBIR system is done by combining two or more areas simultaneously to achieve better results in terms of performance.

9. Conclusions

In this study, various comprehensive approaches are investigated. The discussion of the general framework of the CBIR system is carried out. The important points of extraction from the encompassing survey are summarized. The survey shows that CBIR is still a very interesting area for research. Different work is undergoing in this field to develop or modernize the CBIR system. It lessens the gaps between the low-level and high-level features. Most of the studies are deployed to optimize retrieval performance. Some recent work on the CBIR system is done by combining two or more areas simultaneously to achieve better results in terms of performance.

Table 7. Comparison of feature extraction methods

No.	Feature Extraction Method	Performance Evaluation Parameter	Advantage	Disadvantage
1	Color Histogram, Standard Wavelet	Retrieving Accuracy	It improves the retrieval accuracy.	Insufficient feature set
2	Color Moment, Gabor Filter GVF	Retrieving Efficiency	It creates a vigorous feature set.	elevated semantic gap
3	Color Moment, Gabor Filter, Co-occurrence Matrix	Precision	It reduces the semantic gaps via the RF and SVM	It is time overriding to the label negative.
4	Color Histogram, Tamura, Zenike Moment & Edge	Precision and Recall	Reduce the size of the dataset. All images with correlated features that are similar are retrieved.	Similarity measurement and image retrieval perform two times so it increases calculation.
5	Daubechies Wavelet	Precision, classification accuracy	Reduce the size of the search area. Manage a comprehensive image database.	The feature set not sufficient

Declaration

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article. The authors also declared that this article is original, was prepared in accordance with international publication and research ethics, and ethical committee permission or any special permission is not required.

Author Contributions

The authors H. Koyuncu, M. Dixit and B. Koyuncu carried out the investigation. H. Koyuncu and M. Dixit did the literature review. H. Koyuncu, M. Dixit and B. Koyuncu performed the analysis. M. Dixit performed the data curation. H. Koyuncu and M. Dixit performed the writing original draft preparation. B. Koyuncu performed the writing review, and editing.

Acknowledgment

We gratefully appreciate the contributions received at every stage of the study from Devshri Satyarthi working as a lecturer in Dr. B.R.A. Polytechnic College, Gwalior, India.

Nomenclature

<i>CBIR</i>	: Content-Based Image Retrieval
<i>SIFT</i>	: Scale Invariant Feature Transform
<i>BOW</i>	: Power Bag of Visual Words
<i>QBIC</i>	: Query by Image Content
<i>CBVIR</i>	: Content-Based Visual Information Retrieval
<i>FCM</i>	: Power Fuzzy C-Means

References

- Alzu'bi, A., A. Amira, and N. Ramzan, *Semantic content-based image retrieval: A comprehensive study*. Journal of Visual Communication and Image Representation, 2015. **32**: p. 20-54.
- Li, X., et al., *Socializing the Semantic Gap*. ACM Computing Surveys, 2016. **49**(1): p. 1-39.
- Lin, Z., et al. *Semantics-preserving hashing for cross-view retrieval*. in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2015. IEEE.
- Lowe, D.G., *Distinctive Image Features from Scale-Invariant Keypoints*. International Journal of Computer Vision, 2004. **60**(2): p. 91-110.
- Feng, J., Y. Liu, and L. Wu, *Bag of Visual Words Model with Deep Spatial Features for Geographical Scene Classification*. Computational Intelligence and Neuroscience, 2017. **2017**: p. 5169675.
- An, J., S.H. Lee, and N.I. Cho. *Content-based image retrieval using color features of salient regions*. in *2014 IEEE International Conference on Image Processing (ICIP)*. 2014.
- Philbin, J., et al. *Object retrieval with large vocabularies and fast spatial matching*. in *2007 IEEE Conference on Computer Vision and Pattern Recognition*. 2007.
- Hamouchene, I. and S. Aouat, *A new approach for texture segmentation based on NBP method*. Multimedia Tools and Applications, 2017. **76**(2): p. 1921-1940.
- Sun, S., et al., *Scalable Object Retrieval with Compact Image Representation from Generic Object Regions*. ACM Trans. Multimedia Comput. Commun. Appl., 2015. **12**(2): p. Article 29.
- Jenni, K., S. Mandala, and M.S. Sunar, *Content Based Image Retrieval Using Colour Strings Comparison*. Procedia Computer Science, 2015. **50**: p. 374-379.
- Liu, M., L. Yang, and Y. Liang, *A chroma texture-based method in color image retrieval*.Optik, 2015. **126**(20): p. 2629-2633.
- Batko, M., et al., *Content-based annotation and classification framework: a general multi-purpose approach*, in *Proceedings of the 17th International Database Engineering & Applications Symposium*. 2013, Association for Computing Machinery: Barcelona, Spain. p. 58-67.
- Cheng, Z., J. Shen, and H. Miao, *The effects of multiple query evidences on social image retrieval*. Multimedia Systems, 2016. **22**(4): p. 509-523.
- Zhang, H., et al., *Attribute-Augmented Semantic Hierarchy: Towards a Unified Framework for Content-Based Image Retrieval*. ACM Trans. Multimedia Comput. Commun. Appl., 2014. **11**(1s): p. Article 21.
- Papushoy, A. and A.G. Bors, *Image retrieval based on query by saliency content*. Digital Signal Processing, 2015. **36**: p. 156-173.
- Raveaux, R., J.-C. Burie, and J.-M. Ogier, *Structured representations in a content based image retrieval context*. Journal of Visual Communication and Image Representation, 2013. **24**(8): p. 1252-1268.
- Sokic, E. and S. Konjicija. *Novel fourier descriptor based on complex coordinates shape signature*. in *2014 12th International Workshop on Content-Based Multimedia Indexing (CBMI)*. 2014.
- Bai, C., et al., *K-means based histogram using multiresolution feature vectors for color texture database retrieval*. Multimedia Tools and Applications, 2015. **74**(4): p. 1469-1488.
- Bala, A. and T. Kaur, *Local texton XOR patterns: A new feature descriptor for content-based image retrieval*. Engineering Science and Technology, an International Journal, 2016. **19**(1): p. 101-112.
- Guo, J. and H. Prasetyo, *Content-Based Image Retrieval Using Features Extracted From Halftoning-Based Block Truncation Coding*. IEEE Transactions on Image Processing, 2015. **24**(3): p. 1010-1024.
- Huang, M., et al., *Content-based image retrieval technology using multi-feature fusion*.Optik - International Journal for Light and Electron Optics, 2015. **126**(19): p. 2144-2148.
- Bakar, S.A., M.S. Hitam, and W.N.J.H.W. Yussof. *Content-Based Image Retrieval using SIFT for binary and greyscale images*. in *2013 IEEE International Conference on Signal and Image Processing Applications*. 2013.
- Matsui, Y., K. Aizawa, and Y. Jing. *Sketch2Manga: Sketch-based manga retrieval*. in *2014 IEEE International Conference on Image Processing (ICIP)*. 2014.
- Montazer, G.A. and D. Giveki, *Content based image retrieval system using clustered scale invariant feature transforms*.Optik, 2015. **126**(18): p. 1695-1699.

25. Chahooki, M.A.Z. and N.M. Charkari, *Shape retrieval based on manifold learning by fusion of dissimilarity measures*. IET Image Processing, 2012. **6**, 327-336.
26. Khodaskar, A.A. and S.A. Ladhake. *A novel approach for content based image retrieval in context of combination S C techniques*. in *2015 International Conference on Computer Communication and Informatics (ICCCI)*. 2015.
27. Shrivastava, N. and V. Tyagi, *Content based image retrieval based on relative locations of multiple regions of interest using selective regions matching*. Information Sciences, 2014. **259**: p. 212-224.
28. Seetharaman, K. and M. Kamarasan, *Statistical framework for image retrieval based on multiresolution features and similarity method*. Multimedia Tools and Applications, 2014. **73**(3): p. 1943-1962.
29. Rahimi, M. and M. Ebrahimi Moghaddam, *A content-based image retrieval system based on Color Ton Distribution descriptors*. Signal, Image and Video Processing, 2015. **9**(3): p. 691-704.
30. Yue, J., et al., *Content-based image retrieval using color and texture fused features*. Mathematical and Computer Modelling, 2011. **54**(3): p. 1121-1127.
31. Singh, N., K. Singh, and A.K. Sinha, *A Novel Approach for Content Based Image Retrieval*. Procedia Technology, 2012. **4**: p. 245-250.
32. Alsmadi, M.K., *An efficient similarity measure for content based image retrieval using memetic algorithm*. Egyptian Journal of Basic and Applied Sciences, 2017. **4**(2): p. 112-122.
33. Ayoobkhan, M.U.A., C. Eswaran, and K. Ramakrishnan, *CBIR system based on prediction errors*. Journal of Information Science and Engineering, 2017. **33**(2): p. 347-365.
34. Saeed, M.G., F.L. Malallah, and Z.A. Aljawaryy, *Content-based image retrieval by multi-features extraction and k-means clustering*. International Journal of Electrical, Electronics and Computers, 2017. **2**(3): p. 1-11.
35. Joshi, K.D., S.N. Bhavsar, and R.C. Sanghvi, *Image retrieval system using intuitive descriptors*. Procedia Technology, 2014. **14**: p. 535-542.
36. Sharma, M. and A. Batra, *Analysis of distance measures in content based image retrieval*. Global Journal of Computer Science and Technology, 2014. **14**(2): p. 7.
37. Mistry, Y., D.T. Ingole, and M.D. Ingole, *Content based image retrieval using hybrid features and various distance metric*. Journal of Electrical Systems and Information Technology, 2018. **5**(3): p. 874-888.
38. Prasad, D. and V. Mukherjee, *A novel symbiotic organisms search algorithm for optimal power flow of power system with FACTS devices*. Engineering Science and Technology, an International Journal, 2016. **19**(1): p. 79-89.
39. Pradeep, S. and L. Malliga. *Content based image retrieval and segmentation of medical image database with fuzzy values*. in *International Conference on Information Communication and Embedded Systems (ICICES2014)*. 2014.
40. Katira, C., et al., *Advanced content based image retrieval using multiple feature extraction*. International Journal of Innovative Research in Science, Engineering and Technology, 2015. **4**(10): p. 9805-9812.
41. Gupta, E. and D.R. Kushwah, *Combination of Local, Global and K-Mean using Wavelet Transform for Content Base Image Retrieval*. International Journal of Computer Applications, 2015. **116**(14): p. 5-9.
42. Bagri, N. and P.K. Johari, *A Comparative study on feature extraction using texture and shape for content based image retrieval*. International Journal of Advanced Science and Technology, 2015. **80**: p. 41 - 52.
43. Giveki, D., et al., *A New Content Based Image Retrieval Model Based on Wavelet Transform*. Journal of Computer and Communications, 2015. **Vol.03No.03**: p. 8.
44. Selvi, P., et al., *Novel image retrieval approach in similarity integrated network using combined ranking algorithm*. World Engineering & Applied Sciences Journal 2015. **6**(3): p. 152-156.
45. Sasikala, S. and R.S. Gandhi, *Efficient Content Based Image Retrieval System with Metadata Processing*. International Journal for Innovative Research in Science and Technology, 2015. **1**(10): p. 72-77.
46. Singh, S. and E.R. Rajput, *Content based image retrieval using SVM, NN and KNN classification*. International Journal of Advanced Research in Computer and Communication Engineering, 2015. **4**(5): p. 549-552.
47. Dharani, T. and I.L. Aroquiaraaj. *A survey on content based image retrieval*. in *2013 International Conference on Pattern Recognition, Informatics and Mobile Engineering*. 2013.
48. Smeulders, A.W.M., et al., *Content based image retrieval at the end of the early years*. IEEE Trans. Pattern Anal. Mach. Intell., 2000. **22**(12): p. 1349-1380.
49. Li, Z., et al., *Large-scale retrieval for medical image analytics: A comprehensive review*. Medical Image Analysis, 2018. **43**: p. 66-84.
50. Tarulatha, B., N. Shroff, and M.B. Chaudhary. *VIBGYOR indexing technique for image mining*. in *International Conference on Data Mining and Advanced Computing (SAPIENCE)*. 2016.
51. Liu, Y., et al., *A survey of content based image retrieval with high-level semantics*. Pattern Recognition, 2007. **40**(1): p. 262-282.
52. Siradjuddin, I.A., W.A. Wardana, and M.K. Sophan. *Feature extraction using self-supervised convolutional autoencoder for content based image retrieval*. in *2019 3rd International Conference on Informatics and Computational Sciences (ICICoS)*. 2019.
53. Fadaei, S., A. Rashno, and E. Rashno. *Content based image retrieval speedup*. in *2019 5th Iranian Conference on Signal Processing and Intelligent Systems (ICSPIS)*. 2019.
54. Naoufal, M. and N. M'barek. *content based image retrieval based on color string coding and genetic algorithm*. in *2020 1st International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET)*. 2020.
55. Rudrappa, G. and N. Vijapur. *Cloud classification using K-means clustering and content based image retrieval technique*. in *International Conference on Communication and Signal Processing (ICCS)*. 2020.
56. Barbu, T. *Content based image retrieval using gabor filtering*. in *20th International Workshop on Database and Expert Systems Application*. 2009.
57. Juneja, K., et al. *A survey on recent image indexing and retrieval techniques for low-level feature extraction in cbir systems*. in *IEEE International Conference on*

- Computational Intelligence & Communication Technology*. 2015.
58. Mukane, S.M., S.R. Gengaje, and D.S. Bormane, *A novel scale and rotation invariant texture image retrieval method using fuzzy logic classifier*. Computers & Electrical Engineering, 2014. **40**(8): p. 154-162.
 59. Hole, A.W. and P.L. Ramteke, *Content Based Image Retrieval using Dominant Color and Texture features*. International Journal of Advanced Research in Computer and Communication Engineering, 2015. **4**(10): p. 45-49.
 60. Sadek, S., et al. *Cubic-splines neural network- based system for Image Retrieval*. in *2009 16th IEEE International Conference on Image Processing (ICIP)*. 2009.
 61. Hörster, E., R. Lienhart, and M. Slaney, *Image retrieval on large-scale image databases*, in *Proceedings of the 6th ACM international conference on Image and video retrieval*. 2007: Association for Computing Machinery: Amsterdam, The Netherlands. p. 17–24.
 62. Alghamdi, R.A., M. Taileb, and M. Ameen. *A new multimodal fusion method based on association rules mining for image retrieval*. in *MELECON 17th IEEE Mediterranean Electrotechnical Conference*. 2014.
 63. Wang, B., et al. *Saliency distinguishing and applications to semantics extraction and retrieval of natural image*. in *2010 International Conference on Machine Learning and Cybernetics*. 2010.
 64. Zhang, D., M. Monirul Islam, and G. Lu, *Structural image retrieval using automatic image annotation and region based inverted file*. Journal of Visual Communication and Image Representation, 2013. **24**(7): p. 1087-1098.
 65. Su, J., et al., *efficient relevance feedback for content-based image retrieval by mining user navigation patterns*. IEEE Transactions on Knowledge and Data Engineering, 2011. **23**(3): p. 360-372.
 66. Sun, Y. and B. Bhanu. *Image retrieval with feature selection and relevance feedback*. in *IEEE International Conference on Image Processing*. 2010.
 67. Hui, L., et al. *A relevance feedback system for cbir with long-term learning*. in *International Conference on Multimedia Information Networking and Security*. 2010.
 68. Chatzichristofis, S.A., et al., *Accurate image retrieval based on compact composite descriptors and relevance feedback information*. International Journal of Pattern Recognition and Artificial Intelligence, 2010. **24**(02): p. 207-244.
 69. Wankhede, V.A. and P.S. Mohod. *Content based image retrieval from videos using CBIR and ABIR algorithm*. in *2015 Global Conference on Communication Technologies (GCCT)*. 2015.
 70. Sudhakar, M.S. and K. Bhoopathy Bagan, *An effective biomedical image retrieval framework in a fuzzy feature space employing Phase Congruency and GeoSOM*. Applied Soft Computing, 2014. **22**: p. 492-503.
 71. Mukhopadhyay, S., J.K. Dash, and R. Das Gupta, *Content-based texture image retrieval using fuzzy class membership*. Pattern Recognition Letters, 2013. **34**(6): p. 646-654.
 72. Eissa, Y., et al., *Artificial neural network based model for retrieval of the direct normal, diffuse horizontal and global horizontal irradiances using SEVIRI images*. Solar Energy, 2013. **89**: p. 1-16.
 73. Barrena, M., et al., *QatrisiManager: a general purpose CBIR system*. Machine Vision and Applications, 2015. **26**(4): p. 423-442.
 74. Samanta, S., R.P. Maheshwari, and M. Tripathy, *Directional line edge binary pattern for texture image indexing and retrieval*. in *Proceedings of the International Conference on Advances in Computing, Communications and Informatics*. 2012: Association for Computing Machinery: Chennai, India. p. 745–750.
 75. Chu, L., et al., *Robust spatial consistency graph model for partial duplicate image retrieval*. IEEE Transactions on Multimedia, 2013. **15**(8): p. 1982-1996.
 76. Xu, B., et al., *EMR: A scalable graph-based ranking model for content-based image retrieval*. IEEE Transactions on Knowledge and Data Engineering, 2015. **27**(1): p. 102-114.
 77. GuimarãesPedronette, D.C., J. Almeida, and R. da S. Torres, *A scalable re-ranking method for content-based image retrieval*. Information Sciences, 2014. **265**: p. 91-104.
 78. Yu, J., et al., *Learning to rank using user clicks and visual features for image retrieval*. IEEE Transactions on Cybernetics, 2015. **45**(4): p. 767-779.
 79. Hsiao, K., J. Calder, and A.O. Hero, *Pareto-depth for multiple-query image retrieval*. IEEE Transactions on Image Processing, 2015. **24**(2): p. 583-594.
 80. Tiakas, E., et al., *MSIDX: multi-sort indexing for efficient content-based image search and retrieval*. IEEE Transactions on Multimedia, 2013. **15**(6): p. 1415-1430.
 81. Xiao, Z. and X. Qi, *Complementary relevance feedback-based content-based image retrieval*. Multimedia Tools and Applications, 2014. **73**(3): p. 2157-2177.
 82. Lakshmi, A., M. Nema, and S. Rakshit, *Long term relevance feedback: a probabilistic axis re-weighting update scheme*. IEEE Signal Processing Letters, 2015. **22**(7): p. 852-856.
 83. Kundu, M.K., M. Chowdhury, and S. Rota Bulò, *A graph-based relevance feedback mechanism in content-based image retrieval*. Knowledge-Based Systems, 2015. **73**: p. 254-264.
 84. Irtaza, A., M.A. Jaffar, and M.S. Muhammad, *Content based image retrieval in a web 3.0 environment*. Multimedia Tools and Applications, 2015. **74**(14): p. 5055-5072.
 85. Liu, F., et al., *Intelligent and secure content-based image retrieval for mobile users*. IEEE Access, 2019. **7**: p. 119209-119222.
 86. Ahmed, A., *Implementing relevance feedback for content-based medical image retrieval*. IEEE Access, 2020. **8**: p. 79969-79976.

Appendix

Table A.1. Comparative analysis of CBIR techniques

No.	Author and Year	Technique used for feature extraction/ indexing/ matching / relevance feedback/ database used	Comment
1	Chahooki et al. [25]	Shape-based indexing: contour-based and region-based method. Manifold learning is used for dimension reduction the MPEG-7 is the part B and fish shape dataset	Retrieval accuracy is increased due to the combination of four different characteristics of shape features. Isomap manifold learning method increases the retrieval precision
2	Tiakas et al. [80]	Multi sort indexing (MSINX) is high dimensional image descriptor Image Clef Wikipedia Retrieval 2010, Flickr 1 M, IRISA datasets	The system gives a more accurate retrieval result in less time. Mean average precision is calculated. It can handle energetic operations of insertion and deletion in real-time.
3	Batko et al. [12]	Automatic image annotation and classification(semantic search). Precision and response time of Proof media photo bank and proof media search log Annotation is calculated	Focused on web annotation Combines image and text processing techniques Annotation quality can be significantly improved due to various expansion and reduction techniques.
4	Raveaux et al. [16]	The image is segmented into regions. Graph-based image representation (region adjacency graph) is calculated to show spatial relationships From each region color (color histogram), texture (co-occurrence matrices), and shape (Zernike Moments) features are computed K-means clustering is used for cluster regions Coil-100 dataset is used.	This approach gives good results as compare to the tree-based approach.
5	Chu et al. [75]	Partial Duplicate Image Retrieval (PDIR) using SIFT features. Combined orientation position (COP) consistency Graph model for similarity matching, calculates the mean average precision and Average Retrieval Time. It uses Holidays/1000k, Sub-Dupimage/1000k, Dupimage 1000k, IPDID/1000k, and mobile data set	It enables us to accurately match visual Words. It is between two partial duplicate images. As PDIR is a system factor, it improves the strength of dealing with different data, as it is based on SIFT feature extraction. The method is proved as effective in retrieving near-duplicate images.
6	Samanta et al. [74]	Image indexing and retrieval using Line Edge Binary Pattern (LEBP) Brodatz image dataset Performance is measured using the Average Retrieval Rate (ARR)	DLEBP extracts eight directional line edge values as well as line edge information. Image retrieval performance significantly improves known as Average retrieval rate (ARR)
7	Bakar et al. [22]	Scale-invariant feature transform is based on the feature extraction method. Experimentation was done on the MPEG-7 dataset	Mostly suited and provide outstanding retrieval results for images with many good alternatives to traditional CBIR system as it is invariant to scale rotation and translation
8	An et al. [6]	Descriptor works on color features and finds salient objects. A binary map (spatial distribution of dominant color) roughly describes object shape and relative location. Testing is done on the Corel 1k and Corel 10k dataset.	It provides better retrieval performance according to conventional color-based methods. A binary map of dominant color matches the shapes well, Therefore, it is most suitable for object-based color image retrieval
9	Cheng et al. [13]	Color (color histogram, colorcorrelogram) and texture (Gabor, Tamura, and Edge histogram) visual features textual information (social-tag based) retrieval method relevance feedback NUS-WIDE and MIRFLICKR dataset	Retrieval system using the textual feature can achieve much better performance than only visual features
10	Zhang et al. [14]	Tries to reduce the Semantic gap used Hybrid, feedback mechanism to refine search result	The method can be used to distinguish the semantic gap, the resemblance with image accuracy. user search images quickly.
11	Pedronette et al. [77]	Image re-ranking using BP-tree etc., Result (MAP) evaluated with different feature includes ACC, BIC, CCV, GCH and LCH, ALOI dataset with 72000 images and 1000 classes of objects	The rank list is produced by an efficient indexing structure. It is scalable and well suited to a large dataset
12	Shrivastava et al. [27]	Region-based on segmentation of overriding color and local binary pattern features extraction of each region MPEG7 CCD and Corel image dataset. Average Normalized Modified the Retrieval Rank of employed calculate the performance	ROI is also used to specify the spatial location of regions It improves efficiency through the feature set containing a dominant color and LBP. It also consumes less computation time
13	Sokic et al. [17]	Fourier descriptor-based feature extraction MPEG-7 CE-1 set B, Swedish leaf dataset is used.	The method outperforms both the effectiveness and efficiency, Not suitable for region-based approaches. This descriptor is essentially a contour-based

Table A.1. Comparative analysis of CBIR techniques (continue)

14	Seetharaman et al. [28]	Multi-resolution-based features extraction Vistex texture DB and Brodatz texture image dataset are used. The average precision and recall rate is calculated.	The system is theoretically easy and memory-efficient. It lowers the computing difficulty. It's ideal for image databases with a lot of data.
15	Xiao et al. [81]	Relevance feedback Combines are high-level semantic and low-level visual features. Datasets used are: COREL images, Flickr Images, NUS-WIDE images	Features Datasets used are: COREL images, Flickr Images, NUS-WIDE images
16	Alzu'bi et al. [1]	Semantic image retrieval to reduce the semantic gap is discussed. Various relevant feedback schemes are explained.	It is explained how the system's performance in terms of accuracy and speed is affected.
17	Bai et al. [18]	Color and texture features are used to construct a Multiresolution feature vector. For the classification of the number of histograms, K-means histograms are used. Widely used texture databases are selected: VisTex, A LOT and Stex	Easy implementation Improves retrieval performance compare to state-of-art techniques.
18	Lakshmi et al. [82]	Relevance feedback by axis re-weighting scheme is proposed Caltech and Corel dataset is used for testing	An approach that leads to better convergence reduces the digit of iterations to reach superior retrieval accuracy.
19	Papushoy et al. [15]	Defining salient regions at local and global level Earthmovers distance is used for similarity comparison Benchmark. Dataset used are Simplicity and Corel 1K	It can produce a similar outcome to relevance feedback-based retrieval, and the system provides a stable outcome for a large variety of image categories
20	Hamouchenel et al. [8]	Texture segmentation using neighbors-based binary pattern method Brodatz dataset is used.	Research textures have been well recognized, produces better segmentation results compare to the classical decomposition method Improve accuracy of segmentation.
21	Barrena et al. [73]	Color, texture, and shapes feature extraction Classification using automated learning is used, Indexing and relevance feedback is used to increase the retrieval performance	Three spaces in combination improve results for recall and precision Relevance feedback enhances the quality of the retrieval process. Query finding are listed and sorted
22	Sun et al. [9]	Region identification with generic object detection Fusion of CNN and VLAD features Benchmark dataset: Holidays and UK Bench dataset	Promising accuracy is achieved. The system developed is scalable. Image demonstration of competent to the memory visual projection of the retrieval process gives the time efficiency.
23	Balaet al. [19]	Local text on the XOR patterns features descriptor Corel dataset is used in experimentation.	The feature vector is built using LTxXORPs, and the HSV histogram shows a major increase in recall and precision.
24	Matsui et al. [23]	Fine multi-scale edge orientation histogram-based feature extraction is proposed. Magna dataset is used for comparison	A proposed good solution to sketch-based image retrieval Could retrieves images from the MANGA database (not for other sketches)
25	Khodaskaret al. [26]	CBIR system using Ontology, Tries to reduce Semantic Gap using shared vocabulary (semantic features)	Bridges the semantic gap between the low and high-level features Improves semantic image retrieval with high accuracy, precision, and recall
26	Rahimi et al. [29]	Color ton distribution descriptors based on color co-occurrence matrices Classification using self-organizing map Corel and VisTex dataset	It chooses and extracts appropriate visual features with rich content. The low convolution of the feature extraction method is used in the SOM classifier. It can be used as a structural and signal processing feature description system which fails to supply satisfactory result in an image with the strong color distribution.
27	Kundu et al. [83]	Feature extraction using Multi-scale geometric analysis (MGA) of non-sub sampled contourlet transform (NSTC) Graph-based relevance feedback for ranking simplicity dataset, OLIVA dataset, and Caltech dataset is used for testing purposes.	User representation would process retrieval loop to eliminate the semantic gap by reducing the dimensionality of features. The ranking mechanism successfully uses user input to enhance the retrieval process' quality.
28	Irtaza et al. [84]	CBIR using Genetic Algorithm and SVM. Assures effective retrieval by taking user considerations into an account (i.e. Relevance feedback) Corel set A, Corel set B dataset are used.	Genetically optimized SVM overcomes the limitations of regular SVM like classifier instability, hyperplane bias. Image retrieval results show superiority in terms of recall and precision
29	Xuet al. [76]	Gabor wavelets, Grid color time, local binary pattern, edge histogram, and GIST features were used to extract features. The graph-based ranking is used as experimentation on Corel and MNIST dataset	Supports scalable grap reconstruction Significantly reduces the computational time

Table A.1. Comparative analysis of CBIR techniques (continue)

30	Jenni et al. [10]	Color string coding and string comparison-based feature extraction use SVM as a classifier, Corel photo collection is used.	Decreases computational complexity Significantly increases accuracy in image retrieval
31	Yu et al. [78]	The user's input is used to extract visual characteristics (color, shape, and texture). The dataset is sourced from the Microsoft Bing image search engine.	A more precise and robust ranking model is being developed. Visual contents can reduce noise in clicked features.
32	Montazer et al. [24]	Scale-invariant feature transform (SIFT) is based on the K-means clustering algorithm and its classification is Tested on Caltech 101 dataset	Using k-means clustering two main drawbacks of SIFT i.e. memory usage and matching time are overcome which shows high performance in searching images with objects
33	Guo et al. [20]	Error Diffusion Block Truncation Coding was used to extract features (EDBTC). To create feature vectors (BHF), color histogram feature (CHF) and bit pattern histogram feature (BPHF) are used. Image feature descriptors are computed using vector quantization. For the Corel 1000 and Corel 10000 datasets, the APR is determined.	The method is superior to former BTC methods. Due to the added indexing scheme, it achieves higher retrieval accuracy Feature vector are computed by incorporating vector quantization
34	Huang et al. [21]	Color moment (color), Zernikmoments (Shape), and co-occurrence matrix (texture) based feature extraction on Corel image dataset	A combination of three features solves the shortcoming (i.e. partly express and description of the image) of the method as a single feature Retrieval result is better than Contrary retrieval because it takes more time
35	Liu et al. [11]	Feature extraction based on (Chroma) color (graycolor co-occurrence matrices, Tamura and wavelet) texture features	Gives better performance with traditional luminance texture images
36	Hsiao et al. [79]	EMR (efficient manifold ranking) and the Pareto front procedure are capable of handling multiple queries with queries belonging to various image semantics.	Linear combination of ranking results is suitable for real-world datasets
37	Liu et al. [85]	An IND-CPA secure CBIR framework was suggested and introduced, which allows users to retrieve images from the cloud without having to interact with them constantly.	The CBIR framework is effective and reliable, according to the findings.
38	Ahmed [86]	This study proposes a novel relevance feedback retrieval technique(RFRM) for CBMIR. The feedback is applied using the voting values obtained from each class in the image repository.	The suggested RFRM approach outperforms all others in terms of recall and accuracy.