



Balkan Journal of Electrical & Computer Engineering

An International Peer Reviewed, Referred, Indexed and Open Access Journal

www.bajece.com

Vol : 8
No : 1
Year : 2020
ISSN : 2147 - 284X



It is abstracted and indexed in, Index Google Scholarship, the PSCR, Cross ref, DOAJ, Research Bible, Indian Open Access Journals (OAJ), Institutional Repositories (IR), J-Gate (Informatics India), Ulrich's, International Society of Universal Research in Sciences, DRJI, EyeSource, Cosmos Impact Factor, Cite Factor, SIS Scientific Indexing Service, IJIF, iifactor. ULAKBİM-TR Dizin.

General Publication Director & Editor-in-Chief
Musa Yılmaz, Batman University, Turkey.

Vice Editor
Hamidreza Nazaripouya, University of California Riverside, USA

Scientific Committee
Abhishek Shukla (India)
Abraham Lomi (Indonesia)
Aleksandar Georgiev (Bulgaria)
Arunas Lipnickas (Lithuania)
Audrius Senulis (Lithuania)
Belle R. Upadhyaya (USA)
Brijender Kahanwal (India)
Chandar Kumar Chanda (India)
Daniela Dzhonova-Atanasova (Bulgaria)
Deris Stiawan (Indonesia)
Emel Onal (Turkey)
Emine Ayaz (Turkey)
Enver Hatimi (Kosovo)
Ferhat Sahin (USA)
Gursel Alici (Australia)
Hakan Temeltaş (Turkey)
Ibrahim Akduman (Turkey)
Jan Izykowski (Poland)
Javier Bilbao Landatxe (Spain)
Jelena Dikun (Lithuania)
Karol Kyslan (Slovakia)
Kunihiko Nabeshima (Japan)
Lambros Ekonomou (Greece)
Lazhar Rahmani (Algerie)
Marcel Istrate (Romania)
Marija Eidukeviciute (Lithuania)
Milena Lazarova (Bulgaria)
Muhammad Hadi (Australia)
Muhamed Turkanović (Slovenia)
Mourad Houabes (Algerie)
Murari Mohan Saha (Sweden)
Nick Papanikolaou (Greece)
Okyay Kaynak (Turkey)
Osman Nuri Ucan (Turkey)
Ozgur E. Mustecaplioglu (Turkey)
Padmanaban Sanjeevikumar (India)
Ramazan Caglar (Turkey)
Rumen Popov (Bulgaria)
Tarek Bouktir (Algeria)
Sead Berberovic (Croatia)
Seta Bogosyan (USA)
Savvas G. Vassiliadis (Greece)
Suwarno (Indonesia)
Tulay Adali (USA)
Yogeshwarsing Calleecharan (Mauritius)
YangQuan Chen (USA)
Youcef Soufi (Algeria)

Aim & Scope

The journal publishes original papers in the extensive field of Electrical-Electronics and Computer engineering. It accepts contributions which are fundamental for the development of electrical engineering, computer engineering and its applications, including overlaps to physics. Manuscripts on both theoretical and experimental work are welcome. Review articles and letters to the editors are also included.

Application areas include (but are not limited to): Electrical & Electronics Engineering, Computer Engineering, Software Engineering, Biomedical Engineering, Electrical Power Engineering, Control Engineering, Signal and Image Processing, Communications & Networking, Sensors, Actuators, Remote Sensing, Consumer Electronics, Fiber-Optics, Radar and Sonar Systems, Artificial Intelligence and its applications, Expert Systems, Medical Imaging, Biomedical Analysis and its applications, Computer Vision, Pattern Recognition, Robotics, Industrial Automation.



ISSN: 2147- 284X
Vol: 8
No : 1
Year: July 2020

CONTENTS

- A. Özmen, B. Tander, H. Şenol;** Performance of Cellular Neural Network Based Channel Equalizers.....**1-6**
- E. Dursun, S. Varbak Nese, B. Kilic;** Green building certification of urban public railway transport systems for sustainable cities,**7-15**
- N Gedik;** A New Feature Extraction Approach Using Contourlet Transform and T-Test Statistics for Mammogram Classification,.....**16-20**
- C. Catal, A. Kassahun, H. Jan Hoving;** Improving Farm Management Information Systems with Data Mining,**21-30**
- M. Gökdağ, O. Gülbudak;** Model Predictive Control of an Indirect Matrix Converter with Active Damping Capability,**31-39**
- F. Erken;** The Impact of the Government's Incentives on Increasing Investment in Turkey's Solar Photovoltaic Power Plants,**40-49**
- İ. F. Kılınçer, F. Ertam, A. Şengür;** Automated Fake Access Point Attack Detection and Prevention System with IoT Devices,.....**50-56**
- H. E. Kiziloz;** On Base Station Localization in Wireless Sensor Networks,.....**57-61**
- T. C. Karalar;** Desinging Analog Mixed Signal Circuits Using Graphene Nano Ribbon Field Effect Transistors,.....**62-66**
- C. Rahebi, M. Al-Jumaili;** A New Bat Optimization Algorithm to Solve EPD Problem Solving with Transmission Loss,.....**67-72**
- Z. Doğan, R. Selçuk;** A Diagnosis of Stator Winding Fault Based on Empirical Mode Decomposition in PMSMs,.....**73-80**
- G. Kayhan, E. Ergün;** Medicinal and Aromatic Plants Identification Using Machine Learning Methods,**81-87**
- L. Gökrem, M. S. Can, S. Aydın;** Hexapod Robot Design and Performance Comparison of Fuzzy and PID Control Methods,.....**88-97**
- H. Karakaya, İ. E. Şen;** Phase Changing Material Usage to Increase the Efficiency of Photovoltaic Panels,**98-102**
- S. G. Eraldemir, Ü. Kiliç, M. Keleş, M. E. Demirkol, E. Yildirim, L. Tamam;** Classification of EEG Signals in Depressed Patients,**103-107**
- V. Çetin, S. Özekes, H. S. Varol;** Effects of Digital Filtering on the Classification Performance of Steady-State Visual Evoked Potential Based Brain-Computer Interfaces,.....**108-113**

BALKAN JOURNAL OF ELECTRICAL & COMPUTER ENGINEERING

(An International Peer Reviewed, Indexed and Open Access Journal)

Contact

Batman University
Department of Electrical-Electronics Engineering
Bati Raman Campus Batman-Turkey

Web: <http://dergipark.gov.tr/bajece>
<https://www.bajece.com>
e-mail: bajece@hotmail.com

Performance of Cellular Neural Network Based Channel Equalizers

A. ÖZMEN, B. TANDER, and H. ŞENOL

Abstract—In this paper, a popular dynamic neural network structure called Cellular Neural Network (CNN) is employed as a channel equalizer in digital communications. It is shown that, this nonlinear system is capable of suppressing the effect of intersymbol interference (ISI) and the noise at the channel. The architecture is a small-scaled, simple neural network containing only 25 neurons (cells) with a neighborhood of $r = 2$, thus including only 51 weight coefficients. Furthermore, a special technique called repetitive codes in equalization process is also applied to the mentioned CNN based system to show that the two-dimensional structure of CNN is capable of processing such signals, where performance improvement is observed. Simulations are carried out to compare the proposed structures with minimum mean square error (MMSE) and multilayer perceptron (MLP) based equalizers.

Index Terms—Cellular Neural Networks, channel equalization, MLP equalizer, MMSE equalizer, repetitive codes.

I. INTRODUCTION

IN DIGITAL communication systems, the signal at the receiver will be the linear combination of time delayed and original transmitted signals, as a result of reflections and diffractions at the media. Therefore, the transmitted signal will reach to the receiver with a significant loss, which is called ISI. Typically, the ratio of the total erroneous bits to the total transmitted bits, which is called Bit-Error-Rate (BER) or Symbol-Error-Rate (SER) at the receiver, can be used as a measure. Equalization is the process that compensates the ISI and the impact of noise at the receiver by providing the maximum possible BER.

In literature, various methods are introduced for equalization. Equalizers that are complex and require too much computational power are designed for the applications where precision is needed rather than speed [1]–[6]. On the other hand, less complex however, relatively less accurate methods are also proposed when the speed is taken into account. A structure that models the equalizer with a transversal filter

ATILLA ÖZMEN, is with Department of Electrical-Electronics Engineering, Kadir Has University, Istanbul, Turkey (e-mail: aozmen@khas.edu.tr)

 <https://orcid.org/0000-0002-3868-1927>

BARAN TANDER, is with Department of Mechatronics Engineering, Kadir Has University, Istanbul, Turkey (e-mail: tander@khas.edu.tr)

 <https://orcid.org/0000-0002-9037-4668>

HABİB ŞENOL, is with Department of Computer Engineering, Kadir Has University, Istanbul, Turkey (e-mail: hsenol@khas.edu.tr)

 <https://orcid.org/0000-0001-5724-0839>

Manuscript received January 31, 2019; accepted Nov 11, 2019.
DOI: 10.17694/bajece.519464

approach called zero-forcing (ZF) and the algorithm based on minimizing the mean square error between the equalizer output and the transmitted signal called MMSE are the most popular ones among these [7], [8].

Although these are commonly used linear transversal filters in channel equalization, their Bit-Error-Rates (BER) are not satisfactory. For this reason, alternative methods were developed in literature including Neural Network based architectures [9], [10]. However, even their BERs are better than the conventional techniques, because of their complex structures, they require too much computational power. At this point, CNN can be a good alternative to them with its simple topology. Furthermore, since the outputs of a CNN can take either -1 or +1 values, it is logical to use it in the reconstruction of Binary Phase Shift Keying (BPSK) signals. Formerly, CNN was employed to compute the coefficients of the linear transversal filters [11], somehow in our work, we directly used it as an equalizer itself. In [12], CNN structure uses a template with size of $3 \times 3 (r = 1)$ and also the method is applied for only non-repetitive codes. But in this work, CNN template size is selected as $5 \times 5 (r = 2)$ and the proposed method is applied to non-repetitive codes as well as to repetitive codes.

The paper is organized as follows: Firstly, the channel equalization process is introduced; secondly, a brief theory of the CNNs and how they are employed as an equalizer are presented. The MMSE and CNN channel equalization processes for repetitive codes are given in Section 4 and Section 5 respectively. The performance of the MMSE, MLP and the proposed CNN based structures are compared at the simulations section, both for classical and repetitively coded data. Finally, the advantages and drawbacks of CNN Equalizers are discussed at the conclusion section.

II. CHANNEL EQUALIZATION

At a digital communication system, the transmitted signal is distorted by ISI and noise factors as shown in Fig.1. Here, $s[n]$ is the original transmitted BPSK signal having a value of either -1 or $+1$, $h[n]$ is the transmission channel causing an ISI, $w[n]$ is the Gaussian noise, $y[n]$ is the received signal with the expression given in (1), $\hat{s}[n]$ is the equalized output.

$$y[n] = s[n] * h[n] + w[n]; \quad (1)$$

Where " * " denotes the linear convolution.

BER is a measure of the perfectness of the $\hat{s}[n]$ output signal which can be expressed as follows:

$$BER = \frac{\text{total number of bit errors}}{\text{total number of transmitted bits}} = \frac{N_e}{N_b} \quad (2)$$

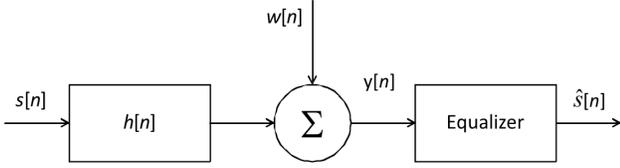


Fig. 1: Model of a digital communication system.

Channel equalization must be carried out in order to reconstruct the original information at the output. Therefore, the motivation of the channel equalization is to design a system that will minimize the difference between the output and the original $s[n]$ signals. In ideal case, making BER=0. Generally, the signal-to-noise-ratio (SNR) versus BER plots are utilized to evaluate the performances of the equalizers for different signal levels.

Although linear transversal filters as well as some special neural network structures, such as MLP can be used for this purpose; in this paper, a system based on CNN is proposed, and the performances are compared for various channel models.

III. CELLULAR NEURAL NETWORKS

CNNs are a class of dynamic neural networks, first proposed by Chua and Yang in 1988 [13] and afterwards, because of their two-dimensional structures, found many impressive applications, especially in image processing [14]. When compared with other neural networks, their demand to less number of weight coefficients also appears to be an advantage, beyond the pros of the two dimensional architecture. As a dynamic

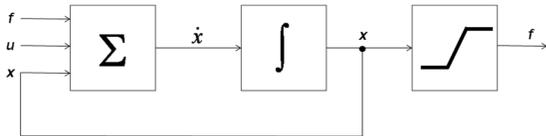


Fig. 2: A CNN cell model.

neural network, a CNN neuron (Cell) seen in Fig.2, which is defined by the partial differential equation in (3), consists of an addition unit, an integration unit and a piecewise-linear activation function.

$$\dot{\mathbf{x}} = -\mathbf{x} + \mathbf{A}f(\mathbf{x}) + \mathbf{B}\mathbf{u} + I \quad (3)$$

Here, \mathbf{x} is the "State", $\dot{\mathbf{x}}$ is its derivative, \mathbf{u} is the input matrix and $f(\mathbf{x})$ is the activation function. \mathbf{A} and \mathbf{B} matrices are called "Cloning" and "Control" templates respectively that assign the interconnection coefficients (weight coefficients) between cells-cells and inputs-cells. Finally, I is a threshold value common for all neurons at the structure.

A stable CNN generates binary outputs $\{-1, 1\}$, since the activation function at the cells defined with (4) is a piecewise-linear function as shown in Fig.3. In other words, a stable CNN is said to approach either to positive or negative saturation regions on the mentioned characteristic. This feature can bring restrictions to many image processing applications however,

from another point of view; the two-level output will be capable of reconstructing the distorted BPSK signals at the receiver end in a digital communication system.

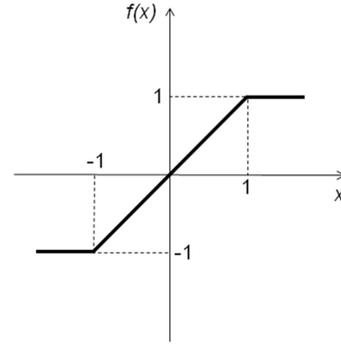


Fig. 3: Activation function of CNN.

$$f(x) = \frac{1}{2} \{ |x+1| - |x-1| \} \quad (4)$$

At an $M \times N$ -cell CNN layer having M cells at its rows and N cells at its columns, the neurons will interact with each other by an r neighbourhood definition given below:

$$N_{ij} = \{ C_{kl} | \max\{|k-i|, |l-j|\} \leq r; 1 \leq k \leq M, 1 \leq l \leq N \} \quad (5)$$

One can see from the above equation that, if $r = 1$, like in most cases, a cell will be only connected to its nearest neighbors, that will dramatically decrease the number of weight coefficients: Specifically, if $r = 2$, the number of coefficients will be 25 for the \mathbf{A} cloning, 25 for the \mathbf{B} control templates which will form 5×5 matrices and an I threshold scalar; totally 51, allowing a relatively simpler implementation.

IV. LINEAR MMSE CHANNEL EQUALIZATION

In a discrete-time communication system the received signal $y[n]$ can be given as follows:

$$y[n] = \sum_{\ell=0}^{N-1} H_{n,\ell} s[\ell] + w[n], \quad (6)$$

Since our system is Linear Time Invariant (LTI), $H_{n,\ell} = h[n-\ell]$. Where $s[n]$ is the input signal, $w[n]$ is additive white Gaussian noise (AWGN) and $H_{n,\ell}$ is the n th row ℓ th column entry of the convolution matrix \mathbf{H} .

Suppose $s[\ell]$ symbols are M -repetitive coded of $d[k]$. By defining $\ell = kM + m$, it is straightforward that $s[kM + m] = d[k]$. So the observation model in (6) can be rewritten as follows:

$$\begin{aligned} y[n] &= \sum_{k=0}^{K-1} \sum_{m=0}^{M-1} H_{n,kM+m} s[kM + m] + w[n] \\ &= \sum_{k=0}^{K-1} \check{H}_{n,k} d[k] + w[n] \end{aligned} \quad (7)$$

where $\check{H}_{n,k} = \sum_{m=0}^{M-1} H_{n,kM+m}$ and this equation actually shows that, \mathbf{H} is separated into K groups each containing M

columns and then $\check{\mathbf{H}}$ is constructed by the summation of the columns in each group. By collecting observation samples, (7) can be represented in vectorial form as follows

$$\mathbf{y} = \check{\mathbf{H}}\mathbf{d} + \mathbf{w} \quad (8)$$

The linear MMSE estimate of the symbols determined from (8) will be as follows [15]

$$\hat{\mathbf{d}} = \left(\check{\mathbf{H}}^\dagger \check{\mathbf{H}} + \sigma_w^2 \mathbf{I}_K \right)^{-1} \check{\mathbf{H}}^\dagger \mathbf{y} \quad (9)$$

However, since $d[k]$ is discrete, belonging to a signal constellation point, we must quantize to its nearest constellation point. Consequently, the detected symbols take the following form

$$\tilde{\mathbf{d}} = \text{Quant}(\hat{\mathbf{d}}), \quad (10)$$

where $\text{Quant}(\cdot)$ denotes the quantization process that quantizes its argument to its nearest data symbol constellation point.

V. CHANNEL EQUALIZATION WITH CNN

CNNs with 25 cells and with a neighborhood of $r = 2$, shown in Fig.4 is employed for the equalization process. Therefore, two 5×5 cloning and control template matrices \mathbf{A} , \mathbf{B} and an I scalar, totally 51 parameters have to be computed as depicted above.

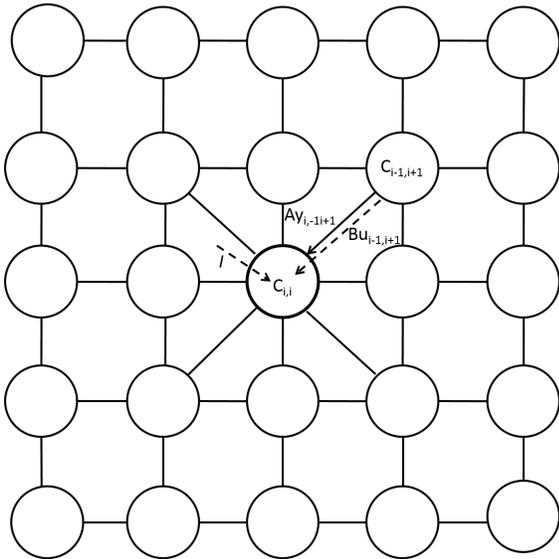


Fig. 4: The 5×5 CNN for the equalization.

A. Classical CNN equalizer design

In order to reorganize the one-dimensional distorted data in two-dimension to work with a CNN, five copies of the distorted data is placed one under the other and the designed CNN is slid above these five rows and a two-dimensional convolution with the 5×5 templates is performed as seen in Fig.5. In this case, each column at the structure will interact between the copies of the $y[k-4]$, $y[k-3]$... $y[k]$ samples at these data rows. Now our goal became the determination of the weight coefficients of the mentioned 5×5 CNN.

Since CNN based equalizer design is to determine the 51 unknown weight coefficients of \mathbf{A} , \mathbf{B} and I , a training process must be carried out as follows: Unequalized data with SNR values varying between 0 and 18dB, employing 1000 symbols for each 1dB range thus, total 19000 samples are applied as the inputs, and the desired uniformly distributed BPSK signals are used as the corresponding outputs for a chosen specific channel. The mentioned unknown templates are found by minimizing the mean square error between the input and desired output with the Genetic Algorithm [16], [17]. After the training, this CNN will be ready to be employed for the new test data.

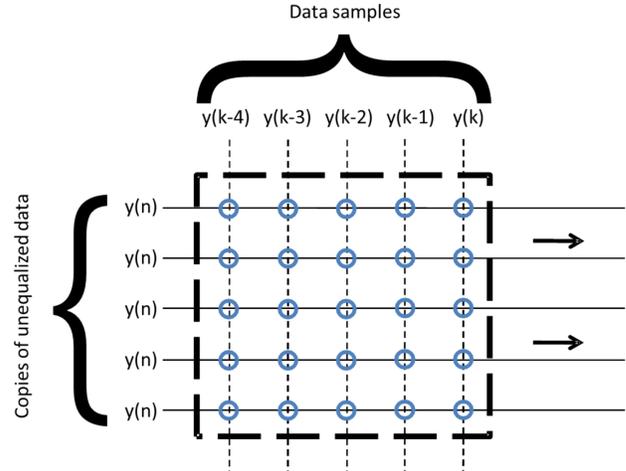


Fig. 5: CNN Equalizer design.

B. CNN Equalizer design with repetitive codes

In this section, the proposed system is applied to repetitively coded data, which means that the transmitted BPSK signal with a chosen length is repeated multiple times consecutively and reorganized as in the classical CNN equalizer design. It is known that, repetition codes provides better BER performance than the classically transmitted data. The detailed process is given below:

- Firstly, the random data is formed as a zero mean BPSK signal with M bits.
- Secondly, the data is converted to a repetitively coded signal by repeating each bit N times, therefore an MN bit sequence is obtained.
- Then, this repetitively coded data is passed through the $h[n]$ channel in Figure 1 while $w[n]$ noise is added.
- Afterwards, before it is applied to the CNN equalizer, the MN bit sequence is paralleled to a 2D structure having N rows and M columns.
- Mentioned $N \times M$ structure is convolved with the CNN equalizer, to obtain stable outputs for each cell.
- If the number of +1s exceeds -1s on a particular column, then the system decides the output as +1 or vice-versa. This decision is indeed the reconstructed BPSK sequence.

All of the steps above at the receiver end are summarized at the block diagram in Fig.6. The training phase of the CNN

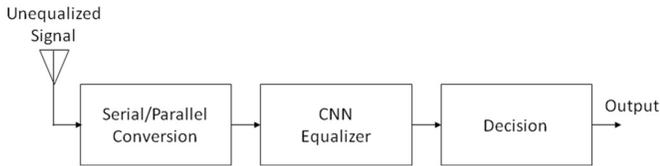


Fig. 6: The receiver for repetition codes

equalizer for repetitive codes, is performed by selecting a transmission signal with various SNRs between -10 and 2 dB in steps of 2 dB and the desired outputs corresponding to these inputs. The data length is chosen as $M = 5000$ for each SNR value. Each bit is repeated three times for 3 repetition codes and five times for 5 repetition codes in order to form a repetitively coded BPSK sequence. Again, the genetic algorithm is employed, where the weight coefficients of the 5×5 **A** and **B** templates are found after.

Afterwards, to observe the performance of the proposed CNN equalizer, BER curves are sketched for data having various SNR values.

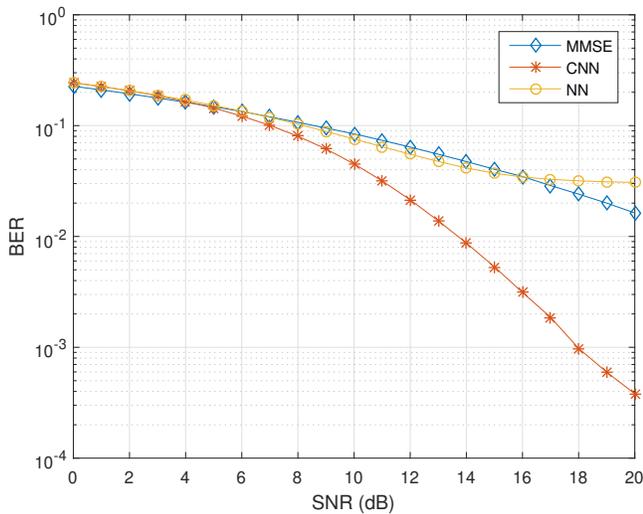


Fig. 7: Comparison of MMSE and CNN BER performances curves for channel B (without repetition).

VI. SIMULATION RESULTS

The CNNs are trained and their BER performances are compared with the MMSE and MLP equalizers for the two individual channels defined with h_B and h_C below with and without repetition codes [7].

$$h_B = \{0.407, 0.815, 0.407\}$$

$$h_C = \{0.227, 0.460, 0.688, 0.460, 0.227\}$$

Simulations are carried out for non-repetition, 3-repetition and 5-repetition cases. After the training phase, the **A**, **B** templates and the I thresholds for the CNNs are computed and then the performance of the MMSE, MLP and CNN equalizers are compared for these cases. Fig.7, Fig.9 and Fig.11 show the

MMSE, MLP and CNN equalizer's BER curves for channel B in the case of non-repetition, 3-repetition and 5-repetition coded signals respectively. Fig.8, Fig.10 and Fig.12 show the BER curves of the equalizers for the channel C at the same cases.

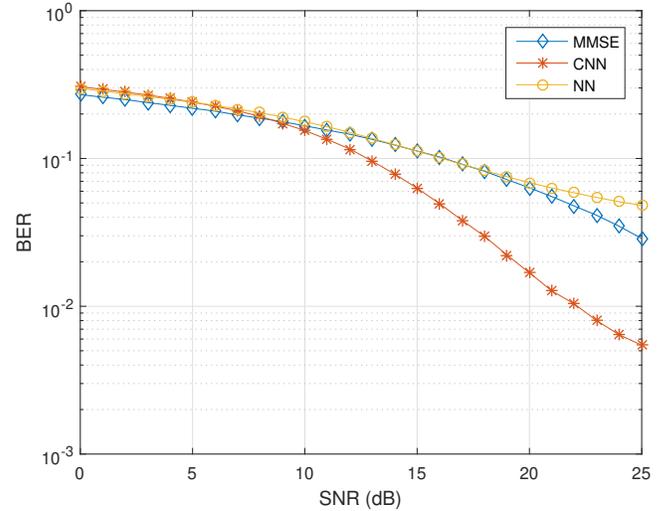


Fig. 8: Comparison of MMSE and CNN BER performances curves for channel C (without repetition).

VII. CONCLUSIONS

In this work, a new channel equalization method employing a CNN is proposed. As CNN's output generates only $+1$ and -1 values, equalization is performed for BPSK signals. Performance of the proposed equalizer is then compared with MMSE and MLP equalizers with and without repetition codes. Furthermore, 3 and 5 repetitions are used for each channel. Performance of the proposed method gives better results specifically at the high SNR values for non-repetition codes. For the 3-repetition coded signals, as shown in the Figures 9 and 10, the performance of CNN equalizer for channel B gives better performance and for channel C gives same performance as MMSE. But, on the other hand, as shown in the Figures 11 and 12, if the number of repetition is increased, the better the performance of MMSE equalizer. However, the matrix that represents the MMSE is quite a large one thus, including too many components therefore, requiring too much computational burden, since the matrix inversion for this purpose consumes too much time and memory when compared with the proposed CNN structure. One can see that, the MLP performance is the worst among aforementioned techniques for all cases.

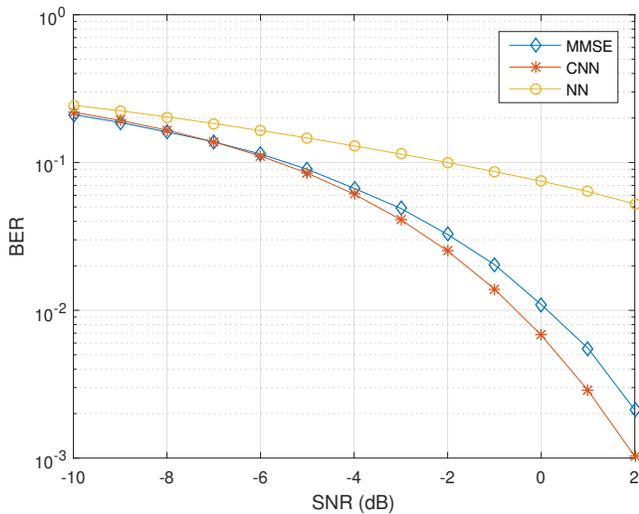


Fig. 9: Comparison of MMSE and CNN BER performances curves for channel B (with 3-repetition).

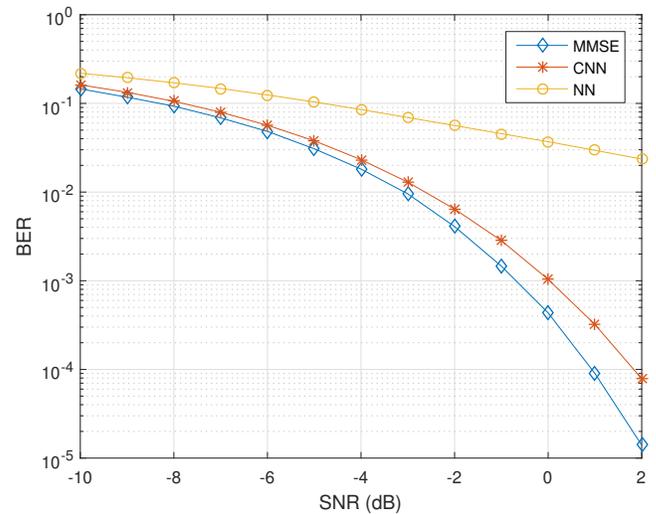


Fig. 11: Comparison of MMSE and CNN BER performances curves for channel B (with 5-repetition).

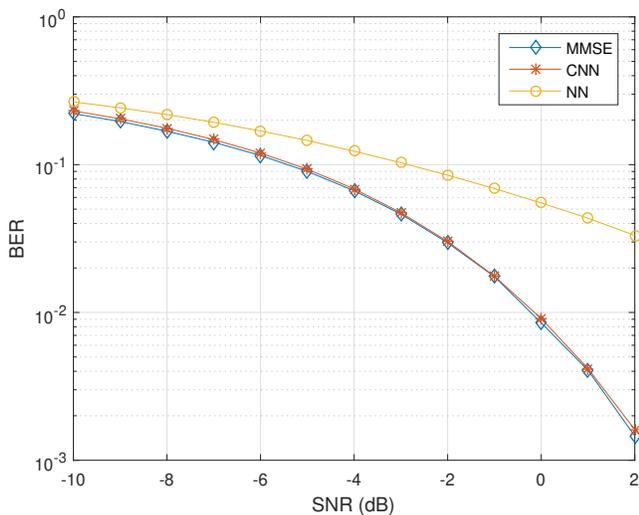


Fig. 10: Comparison of MMSE and CNN BER performances curves for channel C (with 3-repetition).

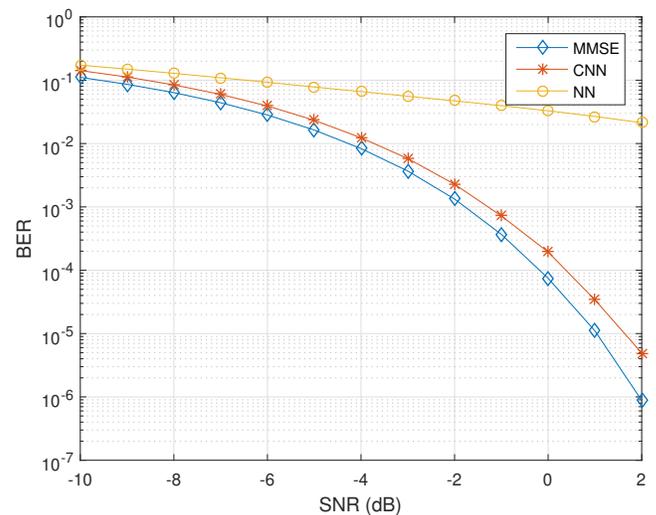


Fig. 12: Comparison of MMSE and CNN BER performances curves for channel C (with 5-repetition).

REFERENCES

- [1] C. Berrou, A. Glavieux, P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding: Turbo-codes," in *Proceeding of IEEE International Conference on Communications*, Geneva, Switzerland, Nov. 30 - Dec. 4 1993, pp. 1064–1070.
- [2] M.K. Lee, K. Yang, "Scheduling for an adaptive number of iterations in turbo equalizers combined with LDPC decoders," *IEEE T Commun*, vol. 58, no. 10, pp. 2759–2764, 2010.
- [3] M. Tuchler, R. Koetter, A. Singer, "Turbo equalization: Principles and new results," *IEEE T Commun*, vol. 50, no. 5, pp. 754–767, 2002.
- [4] S. Talakoub S., L. Sabeti, B. Shahrava, M. Ahmadi, "An improved max-log-MAP algorithm for turbo decoding and turbo equalization," *IEEE T Instrum Meas*, vol. 56, no. 3, pp. 1058–1063, 2007.
- [5] C. He, L. Jing, R. Xi, H. Wang, F. Hua, Q. Dang, Q. Zhang, "Time-Frequency Domain Turbo Equalization for Single-Carrier Underwater Acoustic Communications," *IEEE Access*, vol. 7, pp. 73 324 – 73 335, 2019.
- [6] N. Sharma, J. Sharma, S. Mishra, "Design & implementation of efficient turbo equalizer for noise reduction," in *In: 2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS)*, Chennai, India, August 2017, pp. 2193–2198.
- [7] P. J.G., *Digital Communications 4th ed.* McGraw Hill, 2001.
- [8] B.Gupta, G. Gupta, D. S. Saini, "BER performance improvement in OFDM system with ZFE and MMSE equalizers," in *In: 2011 3rd International Conference on Electronics Computer Technology*, Kanyakumari, India, April 2011, pp. 193–197.
- [9] J. Lee, C. Beach, N. Tepedelenioglu, "A practical radial basis function equalizer," *IEEE T Neural Networ*, vol. 10, no. 2, pp. 450–455, 1999.
- [10] L. Biao, B.L. Evans, "Channel equalization by feed forward neural networks," in *In: Proceeding of IEEE International Symposium on Circuits and Systems*, Orlando, FL, USA, May 30 - June 2 1999, pp. 587–590.
- [11] R. Perfetti, "CNN for fast adaptive equalization," *J Circ Theor*, vol. 21, no. 2, pp. 165–175, March 1993.
- [12] A. Ozmen, B. Tander, "Channel equalization with cellular neural networks," in *In: IEEE Mediterranean Electrotechnical Conference*, Valletta, Malta, April 2010, pp. 1597–1599.
- [13] L.O. Chua, L. Yang, "A Cellular neural networks: Theory," *IEEE Trans CAS*, vol. 35, no. 10, pp. 1257–1272, 1988.

- [14] G. Costantini, D. Casali, M. Carota, "Detection of moving objects in 2-D images based on a CNN algorithm and density based spatial clustering," *WSEAS Trans CAS*, vol. 4, no. 5, pp. 440-447, 2005.
- [15] K. S. M., *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice Hall, 1993.
- [16] T. Tozek, T. Roska, L.O. Chua, "Genetic algorithm for CNN template learning," *IEEE Trans. CAS I: Fund. Theo. and Apps*, vol. 40, no. 6, pp. 392-402, 1993.
- [17] W. Ali, A. A. Ahmed, "Hybrid intelligent phishing website prediction using deep neural networks with genetic algorithm-based feature selection and weighting," *IET Information Security*, vol. 13, no. 6, pp. 659 - 669, 2019.



ATILLA ÖZMEN was born in Batman, Turkey in 1971. He received the B.Sc., M.Sc. and Ph.D. degrees, all in Electronics Engineering, from Istanbul University, Istanbul, Turkey in 1993, 1996 and 2001 respectively. He served as a research assistant at the Department of Electrical and Electronics Engineering, Istanbul University from 1994 to 2001. Currently he is an assistant professor at the Electronics Engineering Department, Kadir Has University. His research interests are neural networks, image processing, signal processing and genetic algorithms.



BARAN TANDER was born in Ankara, Turkey on 1971. He received his BS, MS and PhD degrees in Electronics Engineering all from Istanbul University. He served as a research assistant at the same institution between 1995 and 2001. He is currently an assistant professor at Kadir Has University. His research interests include; neural networks, design and simulation of electronic circuits and image processing and is the author or co-author of more than 20 national and international articles about these subjects.



HABİB ŞENOL (M'87) received the B.S. and M.S. degrees from Istanbul University, Istanbul, Turkey, in 1993 and 1999, respectively, both in Electronics Engineering. He received the Ph.D. degree in Electronics Engineering from Isik University, Istanbul, Turkey, in 2006. He is currently with the faculty of Computer Engineering at Kadir Has University, Istanbul, Turkey. Dr. Senol spent the academic year 2007-2008 at the Department of Electrical Engineering, Arizona State University, USA, working on channel estimation and power optimization algorithms for Wireless Sensor Networks. Dr. Senol's recent research interests include communication theory, advanced signal processing techniques and their applications to wireless electrical, underwater acoustic and optical communications.

Green Building Certification of Urban Public Railway Transport Systems for Sustainable Cities

E. DURSUN, S.VARBAK NESE and B. KILIC

Abstract—The increasing consumption of energy and the greenhouse-gas emissions are challenges in the transport and building sectors. Compared to the other transportation types public railway transport systems are the most energy efficient regarding carbon footprint. Modern urban railway public transport systems have the significant advantages of large passenger capacity, punctuality, safety, going green in the urban environment and reducing carbon footprint, is a scientific consensus. However, the energy demand for railways are increasing in proportion to passenger load, require implementing innovative solutions for optimizing the energy consumption and reducing carbon emission. Leadership in Energy and Environmental Design (LEED) is a building rating system used worldwide. The credit categories of the LEED consist of seven different topics from the Energy and Atmosphere to Materials and Resources. In this study, the LEED certification application to railway station and facilities will be discussed mainly energy efficiency criteria.

Index Terms—Railway buildings, energy efficiency, green building certification, LEED

I. INTRODUCTION

AS THE TRANSPORTATION sector has a significant share in the economy of each country, there is severe competition among the different modes of transport in the European Union countries [1]. Also, urban transport is the source of one-fourth of the CO₂ emissions in the transport sector and 7% of the total greenhouse gas emissions [2]. Road transport has adverse effects such as traffic congestion,

ERKAN DURSUN, is with Department of Electrical and Electronics Engineering of Marmara University, Istanbul, Turkey, (e-mail: erkandursun@marmara.edu.tr).

 <https://orcid.org/0000-0002-7914-8379>

SECIL VARBAK NESE, is with Department of Electrical and Electronics Engineering of Marmara University, Istanbul, Turkey, (e-mail: secil.varbak@marmara.edu.tr).

 <https://orcid.org/0000-0002-1118-5085>

BEYHAN KILIC, is with Department of Electrical and Electronics Engineering of Yildiz Technical University, Istanbul, Turkey, (e-mail: beykilig@yildiz.edu.tr).

 <https://orcid.org/0000-0002-8438-8369>

Manuscript received October 20, 2019; accepted November 24, 2019.
DOI: [10.17694/bajece.649183](https://doi.org/10.17694/bajece.649183)

dependence on fossil fuels, the environment, noise, accident rate [1,2]. For this reason, it is vital to increase the number of integrated, accessible and environmentally friendly transport systems to reduce carbon emissions and improve urban living conditions [2]. Railway transport attracts more attention than road transport. In this context, the European Union countries aim to increase the share of railway transport by 10% until 2020 and as a result, to reduce fuel emissions by 50% [1].

Today transportation is mostly provided by vehicles that use petroleum products as fuel. In railway transportation, it is thought that dependency on petroleum products is decreasing as electric motors are used. On the other hand, in many countries, electricity generation is based on oil and coal-based generation [3,4]. According to the data of 2016, more than 33% of passengers in Japan preferred to use railway transportation, while the number of rail passengers in China was 16.1 billion in urban and 2.8 billion in inter-city. This increased utilization not only increases the infrastructure load, but it also brings with it the requirements of innovation and change [5]. There is also a competitive environment due to the ever-increasing environmental performance and energy costs of other types of transportation, such as the automotive sector. In this competitive environment, it is necessary to reduce energy use while increasing the quality of rail transport services [6].

LEED is the most common international certification system for sustainable buildings with goals such as energy conservation, rational water use, developing a regional project area, improving interior building quality, conscious use of building materials and innovation [7]. Globally, 1.9% of transportation final energy claims and 4.2% of CO₂ emissions from the transport sector in 2015 were due to railway transport [8]. In this context, the U.S. Green Building Council (USGBC) and Green Business Certification Inc. (GBCI) have announced a new LEED pilot rating system for green building to cover transit systems around the world: LEED V4 Operations and Maintenance (O+M): Transit. The new rating system is developed a LEED certification for the design, construction and operation of transport organizations including facilities and stations with a holistic approach. The new performance platform will simplify monitoring the data in railway stations about water, energy usage, waste production, transportation and passenger experience in totally five categories. The LEED management committee, commissioning authority that approves main changes in the technical requirements, validated the new pilot rating system

[9]. In this study, the LEED certification application to railway station and facilities will be discussed mainly energy efficiency criteria.

II. LEED V4 O+M: TRANSIT CERTIFICATION FOR RAILWAY BUILDINGS AND STATIONS

Railway transit systems reduce traffic congestion, emissions, and energy used compared to other transit modes; it improves inherently sustainability of cities. However, the energy demand for railways is increasing in proportion to passenger load, require to implement innovative solutions for optimizing the energy consumption and reducing carbon emission [10]. The USGBC launched a new LEED pilot system to meet the particular requirements of transport systems. This pilot system focuses on the transportation buildings, like passenger stations, depot areas, terminal buildings [11]. Especially railway facilities have high energy usage, water usage, a waste production which affects people lives and environment [12]. The railway station buildings and facilities can take the opportunity to become energy efficient, cost-effective and environmental friendly by implementing LEED v4 O+M: Transit. This pilot system is interpreted as in collaboration with the specific needs of metro stations and facilities LEED certification. It aims at enabling metro stations performance. LEED v4 O+M: Transit can be used to monitor and compare performance across the transit system. The description below and in the certification section outlines how to document performance, identify areas for improvement and certify to LEED. To confirm, project teams must complete two components made up of the LEED v4 O+M rating system [13]. The Arc-online performance platform addresses the need for existing buildings, stations through continuous data uploads, real-time analysis. Even though a building has not certified yet, can participate in the platform, record incremental improvements and performance in five key subcategories. In this context, LEED v4 O+M: Transit will allow transport facilities and stations will enable monitoring and measuring performance, make improvements, score performance and to benchmark efficiencies. This rating system promotes the railway operators reduce their environmental footprint and also create awareness on sustainability. Railway operators can reduce their environmental footprint, while also create awareness on the importance of sustainability and opportunities in minimizing greenhouse gas emissions [14].

Except for the primary structures as rail lines, tunnels, bridges, and tracks, railway systems own to different facilities like railway passenger stations, office buildings, depot areas, maintenance shop, control rooms, traction substations. Inherently, the multi-layered of technical and organizational processes take place at railway facilities, especially railway stations. The challenge is the interdependence of railway

subsystems of complex railway systems to ensure sustainable operation, safety, security and environmental sensitivity. The complex railway systems especially facilities require an integrated application of intelligent technologies and operational strategies to reduce energy consumption and carbon emission. Railway passenger stations are the central component of railway networks, which thousands passenger board and alight complex areas have passenger platforms, concourses, ticket offices. To maintain safe and secure passenger flow in the system is a crucial issue. While organizing the station facilities and infrastructure to ensure a smooth flow of passengers, especially in peak times, energy efficiency, environment-friendly, and sustainable solutions are required [15].

III. THE ACHIEVED LEED CERTIFICATION AT METRO STATIONS AROUND THE WORLD

The LEED v4 O+M: Transit gives an opportunity to address measuring the performance of facilities of metro stations. Railway operators can reduce their energy consumptions, while also drawing the attention to the importance of sustainability in the public transportation sector [16]. Transportation generated 14% of all greenhouse gas emissions in 2010. In the United States, 27% of the greenhouse gas emissions in 2015 were due to transportation. There are more than 160 cities in the world that have developed metro networks, most of which are in Asia and Europe. In a day, more than 160 million passengers travel on metro lines, and only more than 80 million passengers travel on metros in Asia [17].



Fig. 1. Some examples of the LEED-certified railway: Delhi Metro Rail Station(a), Denver Union Station Transit Center(b), The Fulcon Center(c), Rock Island District Station(d) [18-21].

Some examples of the LEED-certified railway station buildings all over the world are shown in Fig.1. The first metro is achieving the pilot, LEED v4 O+M: Transit certification, for Delhi NCR metro stations in India, is Delhi Metro Rail Corporation (DMRC). LEED v4 O+M: Transit will allow transport facilities and stations will allow monitoring and measuring performance, make improvements and score performance and to benchmark efficiencies. The Arc performance platform provides data flow from stations and benchmarking interstations, improvements about five key categories [18]. The Sound Transit's light rail facility has

achieved the LEED Gold certification for green implementations in the construction of the Angle Lake Station in the USA. The station has built environmentally sustainable specifications. 60-solar panel integrated into the platform roof generates energy up to 18 kWh/year. The escalators speed reduces if no passengers. 50 kW solar panels installed on the pedestrian walkway. Four charging stations are installed in the garage. Recycled material that not contains toxic chemicals were used for construction. The station lightings energy consumption reduced by LEDs [22]. The Sound Transit wants to obtain LEED Silver certification for future maintenance facilities [23]. Illinois district metro station has achieved LEED gold certification. The station implemented sustainable applications like PV roof, sensor-fitted lighting [24]. The Metropolitan Transportation Authority has gained LEED Silver rating for the Fulton Center in Manhattan which is the first underground station. Some applications were realized in the main categories as energy efficiency, water usage. More than 25% energy savings were achieved compared to baseline consumption. The Fulton Center provides energy from renewable sources in 50% rate. The other measures are using daylighting reducing electricity consumption, using low-flow plumbing fixtures. The Denver Union Station Transit Center serves local and regional buses, free MetroRide and MallRide buses, commercial Amtrak and commuter light rail. 25% decrease in energy costs and 35% decrease in water usage were observed. Also, 92% construction waste rate has reached the landfill site [25]. Fulton Center on Broadway has LEED Silver certification. The Metropolitan Transportation Authority has achieved 25% energy saved and 40% reduction in water usage with low-flow plumbing fixtures by Fulton Center according to similar constructions. Also, 50% of energy is supplied from renewable sources. In addition to this, electricity usage is decreased using daylight by iconic oculus [26, 27]. Metro downtown Rock Island District Station has LEED gold certification. The solar panel integrated roof, sensor-fitted lighting and the re-use of rainwater systems are used in the station as sustainability features [18]. Canada LEED certification program launched for the building design, construction, and operation processes [23]. GO Transit's Pickering station has gained a first LEED Gold rating and provided 60% energy savings annually. Some of the building's eco-friendly features are listed below: heat pump using, a heat recovery unit, low-flow plumbing fixtures, 45% more potable water, a waste management plan implemented 80% construction waste from landfill. In Germany, the railway hub in Kerpen Horrem supplies its energy through solar panels within the scope of Station Green programme. A geothermal system provides clean and sustainable floor heating at stations [28]. In England, Accrington railway station has a non-heated hall to take advantage of solar gain. Other environmental-friendly features like rainwater harvesting, solar hot water generation and solar photovoltaic technology applied for the energy conservation [29]. In Hong Kong, The Mass Transit Railway Corporation (MTR) serves over seven million people. At the main passenger terminal, 73 escalators and 8 moving walkways were installed when not in use automatically slow down to reduce electricity consumption and, when the built-in radar sensors detect a passenger, it will accelerate. Depending

on passenger volumes, escalators and moving walks with this feature can save up to 60% of the energy consumed. 33 escalators are installed to reduce power consumption and regenerate electricity with the weight of the passengers the steps down, working as a generator [30].

IV. ENERGY EFFICIENCY OPPORTUNITIES FOR RAILWAY TRANSPORT SYSTEMS AND METRO STATIONS

Railways are a vital mode of passenger transportation for both urban and intercity throughout the world. As large-scale public constructions, railway stations are massive structures with high windows and large floor spaces, long platforms, a high ratio of walls, high circulation of people, and have the oversized technical equipment, like escalators, lifts, HVAC systems, lightings. Railway stations also have long operation hours [31]. The oversize of the railway network leads to significant resource consumption. Railway systems' energy consumption can divide into two major categories. The first one is the construction consumptions that consists of train lines, facilities, stations, technical equipment, and trains and after construction process, cannot be changed this consumption. It is important to highlight the decisions during the design process, the construction of infrastructure, trains, and technical systems are hard to change afterward. Therefore, the decision makers should take into considering the environment and energy with a future vision conceptually. The energy consumption during construction of facilities is not covered in this field of research. The second is the whole railway systems energy consumptions that divided into traction of trains and non-traction or auxiliary uses. Traction energy is consumed by running trains while non-traction energy is consumed for the stations, infrastructure facilities, train maintenance workshops, offices. Traction energy is not within the scope of this study. There is an enormous potential for promoting efficiency in non-traction areas of railway facilities, stations, and offices (Table I) [32].

TABLE I
ENERGY SAVING POTENTIAL IN STATIONS AND CONCESSIONS

	Recommended efficiency measure	Subsystem saving potential
Lighting	Intelligent control	10-30%
	Replacement of less efficient lighting	10-40%
Heating and cooling	Intelligent control including temperature setting	10-20%
	Insulation of roof and wall	10-30%
	Renewal of installation	20%
	Modern heat pumps (A+)	Up to 30%
Powered equipment	New efficient drives (A+) and intelligent control	10-40%
Concessions/shops	Targets and energy audits built into contracts	Up to 20% -

In the study [32], railway station buildings including passenger boarding platforms, concourse area. Railway stations can be divided into three main groups: underground, surface and tunnel type. The energy consumptions of the railway stations vary from each other because of different kinds of equipment and features. Divisions of railway systems into subsystems ensure much better the understanding of total and partial energy consumption. Thus, energy consumptions of components or subsystems can be determined. Realizing gains in energy efficiency and emission reduction requires the holistic energy efficiency measures. The system's total energy consumption identified by the division of the systems to subsystems or components.

Railway stations have enormous energy consumption compared with other large public buildings. According to an energy survey, massive stations energy consumption is about 214 (kWh/m²) per year; compared with similar buildings. Accordingly, the station buildings have high-efficiency opportunities for renovating existing buildings. HVAC equipment contributes to 59%–67.9% of all consumption in underground stations [33]. The most significant energy loads were identified as HVAC and lighting systems in underground stations may represent up to 60% of the total consumption [34]. The factors that affect energy consumptions of HVAC systems in the rail facilities should be investigated as temperature setting values and programmable specifications. According to local temperature, HVAC systems' duty cycle must be determined for energy efficiency measures. An energy audit should be accomplished for the lighting and HVAC systems at the facility. Variable Frequency Driven (VFD) HVAC systems can ensure better energy efficiency by 30% [35]. The cooling demand can be reduced by incremental temperature changes from station entrance hall to platform. The exhaust fans can be set according to the outside temperature. If the external environment temperature is lower than 22°C, exhaust fans can blow out the station, air so that cool outer air to be pulled into the station through the entrance. This application reduces energy consumption. The incoming train to the underground station circulates cooled air between the station and the tunnel. The train forces cooled air to the tunnel [36]. This is defined as the piston effect. Platform screen door reduces this effect. Energy consumption of metro systems is shown in Fig.2. Traction consumptions 40%, HVAC is 30%, elevator and lifts consumes 12%. For the lighting system 10%. Finally, 4% for the drainage system [37].

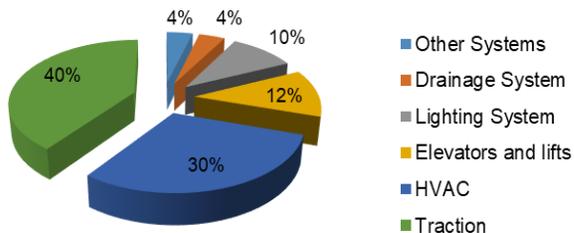


Fig. 2. General energy consumption ratios of metro systems [37].

V. ISTANBUL URBAN RAILWAY SYSTEMS

Istanbul is the most significant city in Turkey which has over 15 million populations. The Metro Istanbul Corporation, the

most prominent operator of urban railways in Turkey, has 174 km of railway lines and transports 1.6 million passengers daily. Electrical energy consumption of The Metro Istanbul Corporation for the 2017 year was 293 million kWh [38]. The on-going new metro line construction's length is 207.6km. In Istanbul urban area, 270.6 km railway lines are in project phase [39]. The Istanbul urban lines and stations specifications are summarized in Table II.

TABLE II
ISTANBUL URBAN RAILWAY LINES SPECIFICATIONS [39]

Line	Name of line	Length (km)	Stations	Escalator number	Lift number	Passenger/day
M1A	Yenikapı-Atatürk Airport	19.75	18	235	65	400000
M1b	Yenikapı-Kirazlı	14.17	13	135	65	400000
M2	Yenikapı-Hacıosman	20.6	16	NA	NA	320000
M3	Kirazlı-Başakşehir	15.9	11	NA	NA	70000
M4	Kadıköy-Tavşantepe	25.6	19	259	70	70000
M5	Üsküdar-Yamanevler	10.6	9	NA	NA	Recently commissioned
M6	Levent-Boğaziçi University	3.1	4	NA	NA	NA
T1	Kabataş-Bağcılar	18.2	31	NA	NA	320000
F1	Taksim-Kabataş	0.594	2	NA	NA	35000
T3	Kadıköy-Moda	2.6	10	NA	NA	1800
T4	Topkapı-Mescid-i Selam	14.5	22	NA	NA	170 000
	Marmaray	13.6	5	NA	NA	75000 /hour

NA: currently not available

M1A light metro line initial started up in 1989. This light metro line consists of partially underground, tunnel, surface, and viaduct stations. The energy consumption of LRT line is 63.35 million kWh in 2015. Since it is the oldest line in Istanbul, the efficiency audit was conducted at one of the tunnel stations. The Bahçelievler station is shown in Fig.3. The station installed power distribution and energy consumption are given in Fig.4 and Fig.5, respectively.

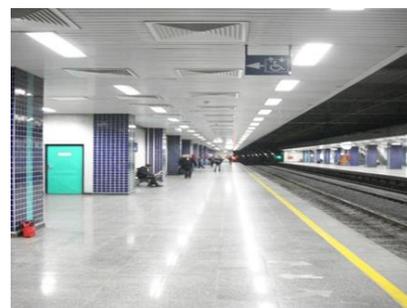


Fig.3. The Bahçelievler tunnel station passenger platform

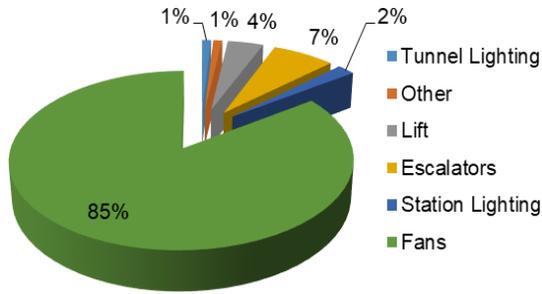


Fig.4. Bahçelievler tunnel station technical equipment installed power (kW) diagram.

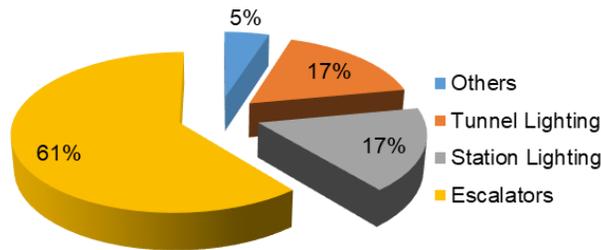


Fig.5. The Atatürk Havalimanı–Yenikapı LRT line Bahçelievler tunnel station monthly energy consumptions [34].

Escalators are one of the most energy consuming systems in underground stations. In the underground stations, the most energy consuming group is HVAC systems [40]. However, in the tunnel type stations, the before and after stations are open stations, HVAC systems are not frequently used because of the piston effect of the trains. The incoming train to the underground station circulates air between the station and the tunnel [36]. Escalators energy consumption can reduce using sensors by 30% compared to non-sensor escalator [41]. The lighting energy consumption takes the second place. Lighting systems ratio to the installed power is 2%. Because of the lighting system is open 24 hours for security and comfort reasons, energy consumptions of lightings systems rise to 34%. The measurement of escalators and platform lightings are given in Fig. 6, 7, 8 and 9 respectively. The escalators in Bahçelievler station have no sensors, Fig.6. The airport station escalators have motion sensors. However, the airport station is out of the scope of this audit, the measurement is shown here for the benchmarking.

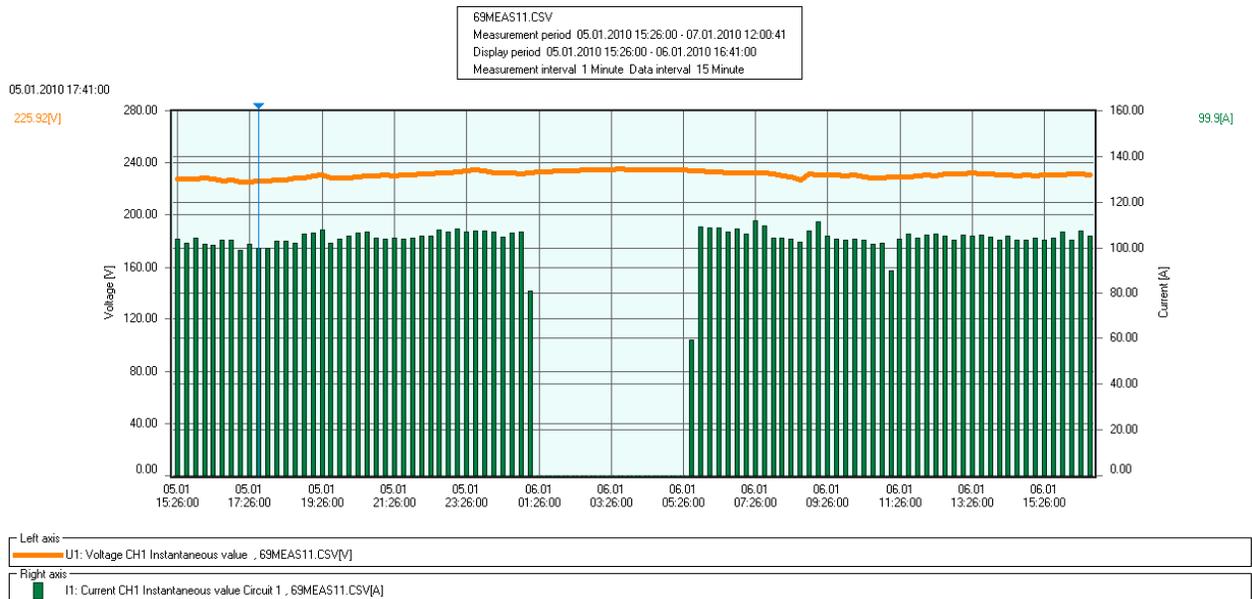


Fig.6. Daily energy consumption of one of the escalators, without motion sensor, at Bahçelievler tunnel station

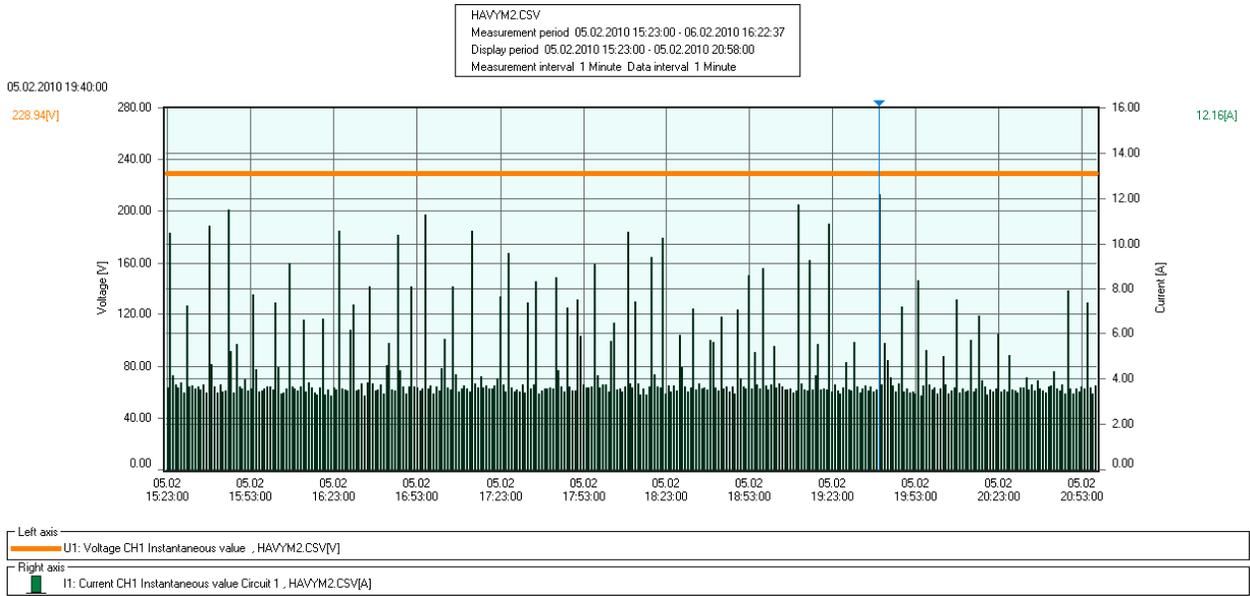


Fig.7. Daily energy consumption of the airport underground station’s escalators with the motion sensor.

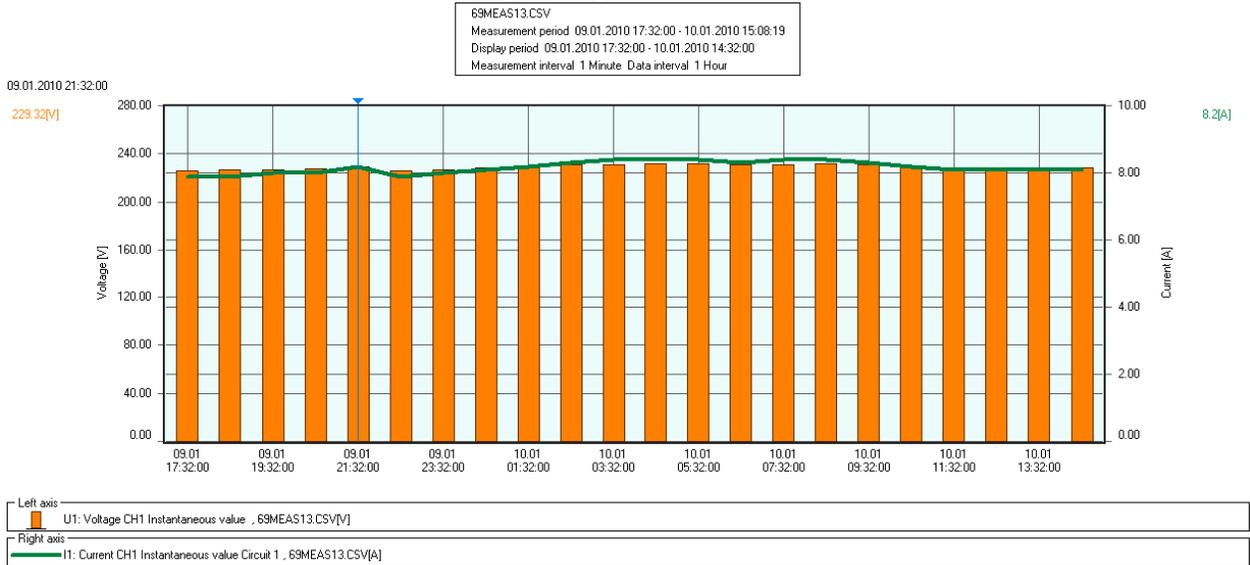


Fig.8. Bahçelievler platform lighting energy consumptions

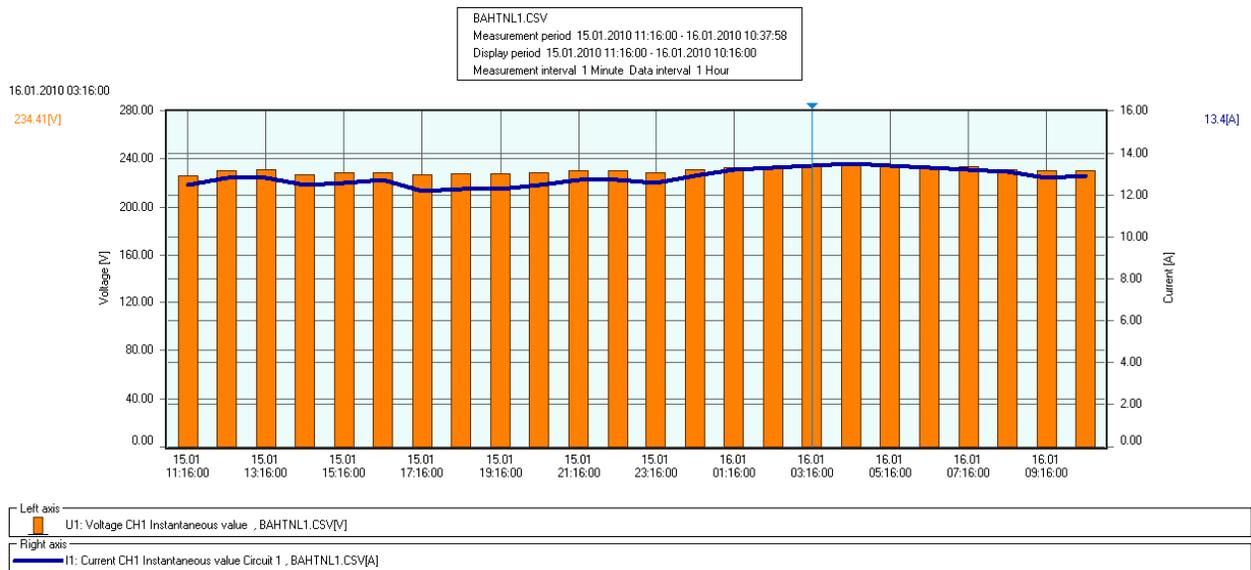


Fig.9. Bahçelievler tunnel lightings energy consumption

Lighting systems at the stations are the second energy consuming category after HVACs. The energy consumptions can be reduced by 28% applying new technologies like LEDs and smart lighting control systems. The general initiatives are identified based on the latest available technologies to reduce the consumption. Energy savings measures ensuring security, safety and comfort for passengers for the existing and elderly metro stations, like Yenikapı-Atatürk Airport LRT line, can be listed below:

-Lighting management in train stations: Railway stations have limited activity on the platforms, tunnels and outdoor spaces at night, especially after 12 pm. The constant light levels make light pollution during the night period. However, the challenge is safety and security problems for passenger and personal when lights off. The energy saving can be provided by using intelligent lighting and sensors in station facilities. The lighting level must be sustained in acceptable minimum values at platforms and tunnels during the night period that trains are at the park area. LED lights usage in stations, information panels, tunnels can provide energy reduction up to 40% and reducing emission for the environment [42].

-Installing renewable sources on-site: The most common renewable energy integration is a solar PV system. The second renewable energy source is wind turbines. Wind turbines in urbanized areas can generate disturbing noise, and the interfering obstacles must be taken away to produce maximum energy. Renewable energy systems must be evaluated observantly regarding availability and sustainability [41]. Renewable sources can partially feed the non-traction loads like HVAC, lifts, escalators, lightings at stations.

-The use of natural light and Led lightings

-High-performance roof for collecting and reusing rainwater: These identified water conservation measures included both demand-side reduction like the use of water-efficient fixtures and supply-side measures like sewage treatment plant and rainwater harvesting. There exists a water saving potential of around 71% [43].

-Installing motion sensors for escalators.

-Using the regenerative braking energy of trains: By the use of the braking energy, there is a reduction of the CO2 emissions. The amount of energy is reduced due to the use of the regenerative brake. There is a reduction potential of approximately 30% of the energy consumption [23].

-Installing platform screen doors

-Policymaking for sustainable improvement: policy-making team and strategists have to be employed by railway operators.

-Building retrofit and energy efficiency projects: Before planning to reduce energy consumption at rail facilities, an energy audit must be executed. Because of significant capital investments of energy saving implementations, a cost analysis requires for determining the payback period.

VI. CONCLUSION

LEED certification is the most common of the certification developed for the standardization of green buildings for a sustainable future. LEED v4 O + M: Transit certification has been introduced for transportation, which has a large share in emissions, waste and energy consumption, such as living areas. It is a well-known fact that railway transport systems are seen as the most innocent especially in terms of environmental factors and so government-promoted means of transportation. On the other hand, greenhouse gas emissions should be minimized in railway transport for sustainability. The energy consumption of railway transport systems due to elevators, escalators, lighting and HVAC systems with long working hours under negative effects results of their physical structures and usage as public areas are essential amounts. In addition, the use of electric motor as an actuator and the fact that the production of electric energy in many countries is based on fossil fuel makes it necessary to provide energy efficiency in railway systems.

REFERENCES

- [1] V. Lingaitis, G. Sinkevičius, "Passenger transport by railway: evaluation of economic and social phenomenon", *Contemp Issues Business Manage Edu, Procedia - Social and Behavioral Sciences*, vol.110, 2014, pp.549–559.
- [2] A. González-Gil, R. Palacin, P. Batty, "Optimal energy management of urban rail systems: Key performance indicators" *Energy Convers Manage*, vol.90, 2015, pp.282–291.
- [3] T. Skrucany, M. Kendra, M. Skorupa, J. Grecnik, T. Figlus, "Comparison of chosen environmental aspects in individual road transport and railway passenger transport", *Procedia Eng* vol.192, 2017, pp.806–811.
- [4] L. Steg, R. Gifford, "Sustainable transportation and quality of life", *J Trans Geogr*, vol.13, 2005, pp.59–69.
- [5] D. He, B. Ai, K. Guan, J.M Garc'ia-Loygorri, L. Tian, Z. Zhong, A. Hrovat, "Influence of Typical Railway Objects in a mmWave Propagation Channel", *IEEE Trans Veh Technol.*, vol.67, 2018, pp.2880–2892.
- [6] A. González-Gil, R. Palacin, P. Batty, J.P. Powell, "A systems approach to reduce urban rail energy consumption", *Energy Convers Manage*. vol.80, 2014, pp.509–524.
- [7] A. Iliopoulos, K. Kirytopoulos, E. Dermitzakis, "Deciding to Go Green: Challenges and Benefits of the LEED Green Building Rating System", 7th International Symposium and 29th National Conference on Operational Research The contribution of Operational Research, new technologies and innovation in agriculture and tourism, Crete, Greece, June, 2018.
- [8] IEA-UIC, *Railway Handbook. Energy Consumption and CO2 emissions focus on passenger rail services*, https://uic.org/IMG/pdf/handbook_iea-uic_2017_web3.pdf [accessed 25 March 2019].
- [9] <https://www.canadianconsultingengineer.com/transportation/leed-system-designed-transit/1003404787> [accessed 25 March 2019].
- [10] D. Pan, L. Zhao, Q. Luo, C. Zhang, Z. Chen, "Study on the Performance Improvement of Urban Rail Transit System", *Energy* vol. 161, 2018, pp.1154–1171.
- [11] LEED in Motion: Transportation, <https://readymag.com/usgbc/transportation> [accessed 29 April 2019].
- [12] Apta Standards Development Program, *Recommended Practice. Transit Sustainability Guidelines, Framework for Approaching Sustainability, and Overview of Best Practices*. American Public Transportation Association, Washington, DC, USA, 2011.
- [13] USGBC Perspective: Understanding Credit Options in LEED v4 O+M, <https://www.facilitiesnet.com/green/article/USGBC-Perspective-Understanding-Credit-Options-in-LEED-v4-OM-Facilities-Management-Green-Feature--17284> [accessed 29 April 2019].
- [14] Green Business Certification, Inc. and the U.S. Green Building Council, *Measure your green performance*, <https://arcskoru.com> [accessed 29 April 2019].
- [15] International Association of Public Transport (UITP), *Shaping the future of our cities and towns: introducing FAIR Stations*, <http://www.uitp.org/news/fairstations> [accessed 27 March 2019].
- [16] U.S. Green Building Council (USGBC), *First LEED Certification for Transit Worldwide Launched*, <https://www.usgbc.org/articles/first-leed-certification-transit-worldwide-launched> [accessed 25 March 2019].
- [17] <http://www.metro-magazine.com/rail/news/724349/sound-transit-station-achieves-leed-gold-certification> [accessed 24 March 2019].
- [18] <http://www.delhimetrorail.com> [accessed 29 April 2019].
- [19] https://www.kawneer.com/kawneer/north_america/en/products/project_detail.asp?project=denver_station&bus=north_america&pc=Remode_l_Retrofit [accessed 27 March 2019].
- [20] <http://www.metro-magazine.com/sustainability/news/711211/mta-s-fulton-center-receives-leed-silver-certification> [accessed 27 March 2019].
- [21] <http://www.metro-magazine.com/sustainability/news/726571/ill-s-metrolink-district-station-earns-leed-gold-certification> [accessed 27 March 2019].
- [22] <https://www.gogreenmetro.com/191/Projects-Planning> [accessed 25 March 2019].
- [23] L. He, Q. Liang, S. Fang, "Challenges and Innovative Solutions in Urban Rail Transit Network Operations and Management: China's Guangzhou Metro Experience", *Urban Rail Transit*, 2016, pp.233–45.
- [24] A. Opeyemi, MD. Cruz, "Green Sustainability Development for Industry Internet of Things in Railway Transportation Industry", *International Conference on Advanced Engineering and Information Technology (ICAET)*, Malaysia, 2017.
- [25] U.S. Green Building Council (USGBC), *What's new in LEED: LEED v4.1.*, <https://www.usgbc.org/articles/whats-new-leed-leed-v41> [accessed 25 March 2019].
- [26] American Public Transportation Association, *Sustainability & Public Transportation Workshop*, San Francisco, CA, 2013. <http://www.apta.com/mc/sustainability/previous/2013/presentations/Presentations/Shatzkin-LEED-Silver-for-Light-Rail-Stations.pdf> [accessed 24 April 2019].
- [27] The Metropolitan Transportation Authority, *Blue Ribbon Commission on Sustainability Renewable Energy Task Report*, <http://web.mta.info/sustainability/pdf/MTA%20Renewable%20Energy%20Report%2010%2029%2008.pdf> [accessed 17 May 2019].
- [28] <http://www.railway-technology.com/features/smart-railway-stations-cities-creating-living-transport-hubs/> [accessed 24 March 2019].
- [29] S. Rzepnicka, D. Załuski, "Innovative_Railway_Stations". *IOP Conf. Series: Materials Science and Engineering* 2017.
- [30] <https://www.smartrailworld.com/9-innovations-that-are-shaping-the-future-of-your-railway-station/> [accessed 27 March 2019].
- [31] <https://newatlas.com/tbox-concept-would-capture-wind-energy-from-speeding-trains/18272/> [accessed 26 March 2019].
- [32] UIC (Union Internationale des Chemins de fer) report, *Study on Non-traction Energy Consumption and Related CO2 emissions from the European Railway Sector*, Copenhagen and Berlin. https://uic.org/IMG/pdf/uic_nontraction_energy_study_final_report_june_2012.pdf [accessed 20 March 2019].
- [33] <https://www.adb.org/sites/default/files/publication/173696/energy-intelligent-railway-station.pdf> [accessed 26 March 2019].
- [34] B. Kılıç, S. Tuna, B. Yagcitezkin, S. Özenc, "Energy efficiency analysis of lighting systems at fixed facilities on light metro systems: a case study". 5th International Ege Energy Symposium and Exhibition (IEESE-5), Turkey, 2010.
- [35] UIC (Union Internationale des Chemins de fer) report (2016). *Technologies and potential developments for energy efficiency and CO2 reductions in rail systems* https://uic.org/IMG/pdf/_27_technologies_and_potential_development_s_for_energy_efficiency_and_co2_reductions_in_rail_systems_uic_in_collaboration.pdf [accessed 24 March 2019].
- [36] P.C. Leung, E.W.M. Lee, "Estimation of electrical power consumption in subway station design by intelligent approach", *Appl Energy*, vol.101, 2013, pp.634–643.
- [37] S. Su, T. Tang, Y. Wang, "Evaluation of strategies to reducing traction energy consumption of metro systems using an optimal train control simulation model", *Energies*, vol.9, 2016, pp.105–124.
- [38] The metro Istanbul annual activity report, http://www.ibb.gov.tr/en-US/Documents/annual_reports/IMM_ANNUAL%20REPORT_2015.pdf [accessed 24 April 2019].
- [39] Istanbul Metropolitan Municipality, www.metroistanbul/ [accessed 27 March 2019].
- [40] J. Liu, B. Lin, Y. Zhang, Q. Zhu, Y. Zhu, "Case study – how could we optimize the energy-efficient design for an extra-large railway station with a comprehensive simulation? Building Simulation", 2011-12th Conference of International Building Performance Simulation Association, Sydney, Australia, 2011.
- [41] K. Fu, Z. Deng, "Current situation of energy consumption in Guangzhou railway station and analysis on the potential of energy conservation", *J Sustain Dev*, vol.2, 2009, pp.117–120.
- [42] *Energy efficiency and reliability solutions for rail operations and facilities (2014). A Report*; <https://www.ctcase.org/reports/railenergy.pdf/> [accessed 29 April 2019].
- [43] Suzuki T. *A Vision of future railway station*, http://www.ejrcf.or.jp/jrtr/jrtr06/pdf/f06_suz.pdf [accessed 29 April 2019].

BIOGRAPHIES



ERKAN DURSUN was born in 1977 in Kayseri, Turkey. He received PhD from Marmara University. His PhD thesis about Distributed Generation (DG), which consist of the photovoltaic/wind/fuel cell hybrid power system. The quality of PhD thesis results is demonstrated by the fact that a paper describing the core contribution of the thesis has been accepted to many scientific journals and conferences. His papers have received more than 400 citations in SCI database of Thomson Reuters. He has been working on power system, renewable energy systems and smart grid for 12 years. He worked as a researcher at European Commission, JRC Institute for Energy and Transport (IET) from 2012 to 2013. He became an Assistant Professor at Marmara University, Faculty of Technology, Department of Electrical and Electronics Engineering in 2013. He is a member of IEEE.



SECİL VARBAK NEŞE received the B.S. degree in Electrical Education from Kocaeli University, Kocaeli, Turkey, in 2004, the M.S. degree in Electrical Education from the Afyon Kocatepe University, Afyonkarahisar, Turkey, in 2008, and the Ph.D. degree in Electrical Education from the Marmara University, Istanbul, Turkey, in 2015. She is currently an Assistant Professor Dr. at Electrical-Electronics Engineering in Marmara University, Istanbul, Turkey. Her current research interests include power systems, renewable energy.



BEYHAN KILIC was born in Yozgat, Turkey. She completed B.Sc. at the Department of Electrical Engineering, Yildiz Technical University and Ph.D. at the same department of Yildiz Technical University respectively in 1985 and 1998. She worked at Istanbul Metropolitan Municipality Energy Department 2007-2009. She was assigned to Istanbul Metro Co. for three years. She is currently as Research Assistant at the Electrical Engineering Department of Yildiz Technical University. Her research areas are: railway systems, integration of renewable energy to railway systems, energy management and smart grid applications in railway systems.

A New Feature Extraction Approach Using Contourlet Transform and T-Test Statistics for Mammogram Classification

N. GEDİK

Abstract— In this study, a CAD system is recommended for the classification of mammographic images. The images are classified as normal-abnormal and benign-malignant. The proposed system consists of three basic steps: the feature extraction, determination of the distinguishing capabilities of the features and selection, and classification. The distinguishing capabilities of the features mean determining the best or optimal features. Thanks to this determination, mammograms could be put into classes with high accuracy. The determination process is carried out using thresholding and t-test statistics. Classification is performed repeatedly for all threshold values using support vector machine. Among the obtained results of the classification, the optimal feature set, which has the best classification performance, is selected. Finally, to evaluate the optimal feature set, classification carries out applying 5-fold cross-validation.

Index Terms— Classification, CAD, Contourlet transform, Mammogram, SVM.

I. INTRODUCTION

CANCER IS one of the most serious health problems globally. Of all the types of cancer, breast cancer is the most common type of cancer and threatens women (especially those over 40 years of age). It is stated that the number of new cases diagnosed as breast cancer is approximately 250,000 each year [1]. Struggling with cancer successfully, and reducing the mortality rates can only be achieved with early diagnosis and appropriate treatment [2, 3]. In order to diagnose breast cancer at an early stage, mammography is the widely used and accepted methodology by radiologists [4]. Screening mammograms is a challenging and busy task, and some cases can be overlooked and misinterpreted. Studies show that 10%-30% of cases of cancer have been overlooked (false negative), and 20%-30% of cases sent to biopsy do not have cancer (false positives) [5, 6]. Development of automated systems is a requirement to support experts as the second opinion in reducing false negative and positive rates. Computer-aided

diagnosis (CAD) systems provide appropriate vision and orientation for diagnosis [7, 8].

The development of CAD systems is a progressive area and one of the most important stages in the system is the selection and extraction of features. There are different approaches to feature extraction in CAD studies [9, 10]. Spatial data can be used to obtain features directly, or a transform can be used to obtain the features via a different domain by transferring the data [11,12]. Gedik [13] proposes a feature extraction approach combining fast finite shearlet transform (FFST) and t-test statistics. FFST is employed to decompose mammograms and to get shearlet coefficients for building the feature matrix. To eliminate the redundant features, a feature selection process is performed using distinguishing capability values of features counting by t-test statistics and thresholding. Jadoon et al. [14] present a mammogram classification system consisting of discrete curvelet transform and multilayer perceptron algorithm. First of all, the images are filtered, and then the curvelet transform is applied to the filtered images. The dense scale invariant feature transform (DSIFT) is applied to the curvelet transform sub-bands to obtain scale-invariant features. Classification is carried out by using multilayer perceptron classifier. Chen et al. [15] present a system that uses wavelet energy entropy and linear regression classifiers to classify mammographic images for breast cancer detection. In the method, the images are segmented, and the regions of interests (ROIs) are obtained. Wavelet energy entropy is calculated applying the wavelet transform to the ROIs. Wavelet energy entropy values are used as features, and the classification process is performed with linear regression classifier. Eltoukhy et al. [16] compare wavelet and curvelet transform to classify the mammograms. Wavelet and curvelet transforms are applied to the images individually, and the feature matrixes are formed by using their coefficients. The features are listed depending on their ability of discriminating the classes using the t-test statistics. Eltoukhy et al. [17], in another study, also use wavelet and curvelet transform for feature extraction comparatively. The features are also sorted according to their ability to distinguish the classes, and then thresholding is applied to

Nebi GEDİK, is with Institute of Health Sciences, University of Health Sciences, Istanbul, Turkey, (e-mail nebi.gedik@sbu.edu.tr).

 <https://orcid.org/0000-0002-1560-1058>

Manuscript received April 24, 2020; accepted January 25, 2020.

DOI: [10.17694/bajece.557693](https://doi.org/10.17694/bajece.557693)

obtain the most effective features. In contrast to the previous study, the difference of class averages is used to determine the ability of features to distinguish the classes. The method proposed by Sehrawat et al. [18] for tumor detection includes wavelet transform and support vector machines. Mammographic images are preprocessed using Gaussian filter, and then the wavelet transform is applied to the images. Features are composed using wavelet coefficients. Classification is carried out by support vector machines.

In this study, an alternative feature extraction and selection method for the classification of mammograms by using modified contourlet transform (contourlet transform with sharp frequency localization, CTsFL) is presented. Initially, the images are decomposed and transformed into coefficients by CTsFL transform, then these coefficients are used to generate the feature matrix (MxN). In order to determine the most effective features, the ability of features to distinguish classes is determined by coherence coefficients using t-test statistics. By applying dynamic thresholding over coherence coefficients, the most effective feature set is assigned. The most effective features are the ones that provide maximum classification accuracy with the minimum number of features. To evaluate the most effective features, the classification process is performed using 5-fold cross-validation with these features.

II. CONTOURLET TRANSFORM WITH SHARP FREQUENCY LOCALIZATION

The method presented by Lu and Do in 2006 [19] improves the frequency domain localization, which is seen as the major disadvantage of the original contourlet transform. While the original contourlet transform uses Laplacian pyramid, Lu and Do use a new multi-scale decomposition defined in the frequency domain. The new pyramid structure generated by Lu and Do is shown in Figure 1.

In the block diagram, $L_i(\omega)$ $i=0,1$ represents the low-pass filters and $D_i(\omega)$ $i=0,1$ represents the high-pass filters. As a significant difference from the Laplacian pyramid, it can be used a different set of low-pass and high-pass filters with the new structure for the first level and all other levels. The low-pass filter $L_i(\omega)$ $i=0,1$ is represented in frequency domain as $L_i(\omega) = L_i^{1d}(\omega_1) \cdot L_i^{1d}(\omega_2)$, where $L_i^{1d}(\omega_1)$ is a 1-D low-pass filter with pass-band frequency $\omega_{p,i}$ and stop-band frequency $\omega_{s,i}$. A smooth transition band is defined as Eq.(1) [19].

$$L_i^{1d}(\omega) = \begin{cases} 1 & \text{for } |\omega| \leq \omega_{p,i} \\ \frac{1}{2} + \frac{1}{2} \cos \frac{(|\omega| - \omega_{p,i})\pi}{\omega_{s,i} - \omega_{p,i}} & \text{for } \omega_{p,i} < |\omega| < \omega_{s,i} \\ 0 & \text{for } \omega_{s,i} \leq |\omega| \leq \pi \end{cases} \quad (1)$$

where $|\omega| \leq \pi$ and, $i=0,1$. The perfect reconstruction condition can be simplified as Eq. (2) for the multiscale pyramid [19]

$$|L_i(\omega)|^2 + |D_i(\omega)|^2 \equiv 1 \quad \text{for } i = 0,1 \quad (2)$$

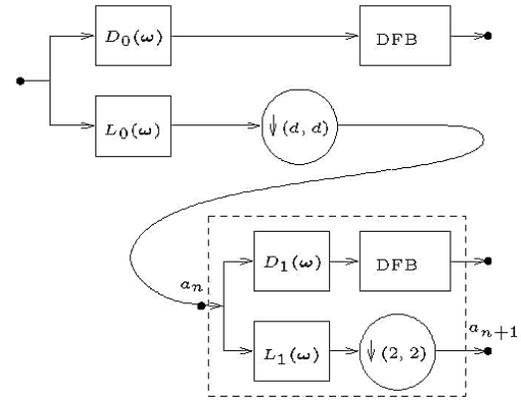


Fig. 1. The new pyramid structure generated by Lu and Do [19].

III. SUPPORT VECTOR MACHINE

SVM is a supervised machine learning algorithm developed by Vapnik for the solution of classification problems [20, 21]. The algorithm is basically based on the Vapnik-Chervonenkis (VC) theory (statistical learning theory) [22]. It translates the input space to a higher dimensional feature space using selected nonlinear mapping. In this feature space, it creates a differential hyperplane and aims to maximize the distance of the closest vectors of different classes to the hyper plane. A training data containing two classes with N feature vectors are defined as follows.

$$X = \{x_1, x_2, \dots, x_N\} \quad (3)$$

The equations for the hyperplane that will separate these feature vectors are described as follows.

$$\begin{aligned} \omega \cdot x_i + b &\geq +1 \\ \omega \cdot x_i + b &\leq -1 \end{aligned} \quad (4)$$

where b the threshold parameter of the hyperplane, ω is the normal of the hyperplane. The distance between the two planes is $\frac{1}{\|\omega\|}$. In order to separate these two classes better, maximum distance must be provided. For the maximum distance, $\|\omega\|$ must be at a minimum value. This requires the solution of the $\min \left[\frac{1}{2} \|\omega\|^2 \right]$ optimization problem using the Lagrange function [20]. The applied Lagrange function gives the most suitable hyper plane and is expressed as follows.

$$L(\omega, b, \alpha) = \frac{1}{2} \|\omega\|^2 - \sum_{i=1}^k \alpha_i y_i (\omega \cdot x_i + b) + \sum_{i=1}^k \alpha_i \quad (5)$$

where α_i are the Lagrange multipliers of the optimization problem. The resulting decision function is as follows.

$$f(x) = \text{sign} \left(\sum_{i=1}^k \alpha_i y_i (\omega \cdot x_i + b) \right) \quad (6)$$

IV. PROPOSED METHOD

Initially, mammograms are decomposed using CTsFL transform and the feature matrix (MxN) is generated from the CTsFL coefficients. Where M (lines) corresponds to every image and N (columns) correspond to CTsFL coefficients (features).

The proposed method aims to identify those which have the best classification performance among the whole feature set. The flow chart of the proposed method is indicated in the Figure 2 step by step. In order to accomplish the objective, t-test statistics by using Eq. (7) [23] is applied to the feature set, and coherence coefficients corresponding to each feature (each column in feature matrix (N)) are obtained. Finally, the coherence coefficients (cc) vector is created.

$$cc = \frac{|\mu_1 - \mu_2|}{\sqrt{\frac{(\delta_1)^2}{N_1} + \frac{(\delta_2)^2}{N_2}}} \quad (7)$$

where cc is the coherence coefficient, $\mu_{1,2}$ and $\delta_{1,2}$ are the mean and the standard deviations of the classes (subscript 1,2 donates normal-abnormal or benign-malignant), and N is the number of images in each class.

By applying dynamic thresholding over cc values, the feature matrix is reconstructed by new dimension (MxD), and the classification is performed using it to find the most effective feature set, iteratively. The generated new feature set is formed from the values of columns of the initial feature matrix (MxN) correspond to cc values greater than the applied threshold. The operations are repeated for all threshold values, and then the optimum feature set, which has the minimum number of feature with maximum accuracy, is determined. To evaluate the optimum feature set, classification is carried out again using 5-fold cross-validation with those features. All classification processes are performed using support vector machine (SVM).

V. EXPERIMENTAL RESULTS

In the evaluation of the proposed method, mammograms obtained from MIAS database [24] are used. Totally 322 images, which consist of 207 normal, 64 benign and 51 malignant, obtained from 161 patients are examined and labeled by expert radiologists. The images are originally 1024x1024 pixels size and about 50% of an image is background and noise. In this paper, those unwanted parts of the images are eliminated by using manual cropping and the ROI set is created. ROIs are 128x128 pixels size, and in case of cropping of abnormal images, the center of ROIs are the center of abnormality. For normal images, the process is randomly made. The ROI set is composed of 228 images including 114 abnormal and 114 normal. In the experiment process, the ROI set is divided into two sets; training (70%) and testing (30%). The training set is used to train the classifier (SVM) and the system performance is evaluated by using the test set. These are illustrated in Figure 2. The image distribution of ROIs is shown in Table 1.

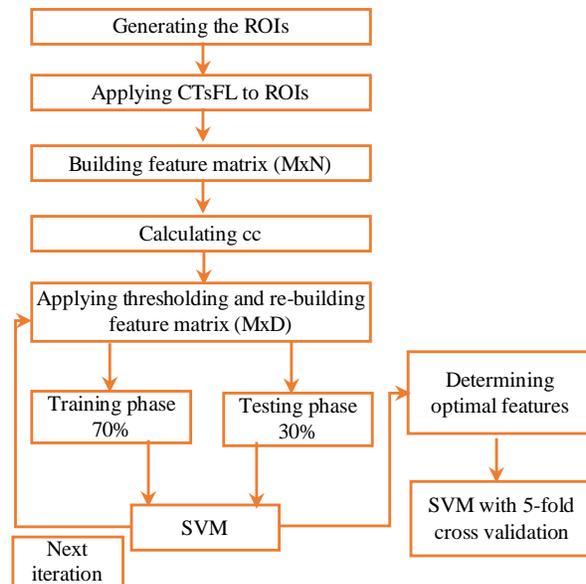


Fig. 2. The flow chart of the proposed method.

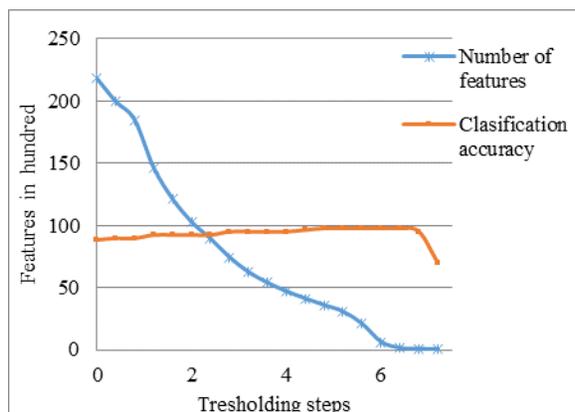
TABLE I
NUMERICAL DISTRIBUTION OF ROIs.

	Benign	Malignant	Total
Abnormal	64	50	114
Normal	—	—	114
Total			228

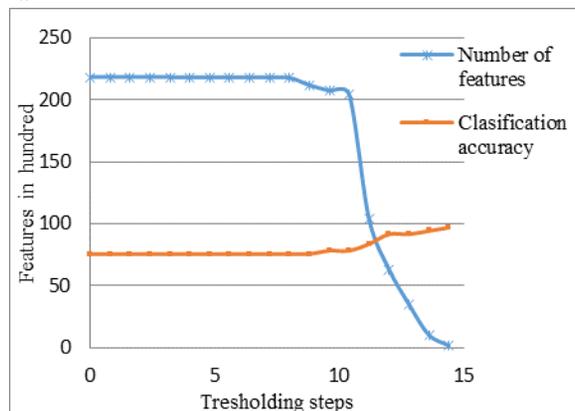
Classification is performed in two-phases; normal-abnormal separation and benign-malignant separation. Figure 3a and 3b illustrate results of the normal-abnormal and benign-malignant classification according to the number of features and the thresholding steps. From the classification results made for all threshold values, maximum accuracies are 97.36% and 97.29% for normal-abnormal and benign-malignant classification respectively. For the maximum accuracies points, while the number of features in the normal-abnormal classification reduces from 21842 to 129, it reduces from 21842 to 175 for benign-malignant classification. Because the goal is to get the maximum accuracy result with the minimum number of features, those points are selected as the optimum features point. To evaluate the result according to the optimum points, classification is performed using 5-fold cross-validation. Figure 4 presents the average value of normal-abnormal and benign-malignant classification results via 5-fold cross validation. Standard deviation values of different classification results obtained from different folds are shown with an error bar above the average classification accuracy bar.

Comparison of the proposed method with previous studies is shown in Table 2. The table contains closely related studies and contains the following comparisons: classification accuracy, number of feature data ensuring the best classification accuracy, feature extraction method, database and classifier. With a general evaluation on the accuracy values in the table, it is observed that the results are close to each other except for [17]. If the normal-abnormal and benign-malignant classification is

evaluated separately, the presented study has the highest accuracy value for the normal-abnormal classification. For benign-malignant classification, the results are almost the same value, except for [17] and for [13] with 98.8% accuracy value. If an evaluation is made in terms of the size of the feature data that used, because all the studies aim to obtain the most successful and the smallest size of feature data from the large feature data, the presented study shows the most successful result.



a



b

Fig. 3. Classifications results a) normal-abnormal separation b) benign-malignant separation.

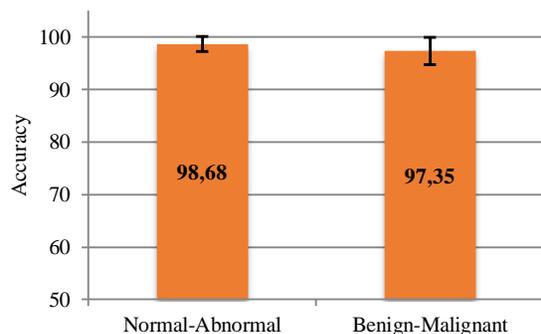


Fig. 4. Average classification results obtained from 5-fold cross validation.

VI. CONCLUSION

In the present study, a CAD system is represented via a new feature extraction approach to classify mammogram images, and its implementation is carried out. The proposed new feature extraction approach includes the contourlet transform and the t-test statistic. The main purpose is to purify the features that make the most contribution to a successful classification by eliminating redundant ones. Success criterion is to achieve maximum classification success by using the minimum number of features among the all obtained features. According to the experimental results, maximum classification rate is 98.68% for normal-abnormal classification using 129 features among 21842 features, while maximum classification rate is 97.35% for normal-abnormal classification using 175 features among 21842 features. Considering the results, the method has an acceptable success to classify mammogram images.

REFERENCES

- [1] M.D. Chin, K.K. Evans, J.M. Wolfe, J. Bowen, J.W. Tanaka, "Inversion effects in the expert classification of mammograms and faces", *Cognitive Research: Principles and Implications*, vol. 3, 2018, pp. 31.
- [2] Y. Wang, H. Shi, S. M., "A new approach to the detection of lesions in mammography using fuzzy clustering", *J. Int. Med. Res.* vol. 39, no. 6, 2011, pp. 2256–2263.
- [3] N.J. Massat, A. Dibden, D. Parmar, J. Cuzick, P.D. Sasieni, S.W. Duffy, "Impact of screening on breast cancer mortality: the UK program 20 years on", *Cancer Epidemiology and Prevention Biomarkers*, vol. 25, no. 3, 2016, pp. 455-62.
- [4] T. Onega, L.E. Goldman, R.L. Walker, D.L. Miglioretti, D.S. Buist, S. Taplin, B.M. Geller, D.A. Hill, R. Smith-Bindman, "Facility mammography volume in relation to breast cancer screening outcomes", *J. Med. Screen.* vol. 23, 2016, pp. 31.
- [5] M.M. Pawar, S.N. Talbar, "Genetic fuzzy system (GFS) based wavelet co-occurrence feature selection in mammogram classification for breast cancer diagnosis" *Perspectives in Science*, vol.8, 2016, pp. 247–250.
- [6] L. Berlin, "Radiologic errors, past, present and future", *Diagnosis*, vol. 1, no. 1, 2014, pp. 79–84.
- [7] Y. Li, H. Chen, Y. Yang, L. Cheng, L. Cao, "A bilateral analysis scheme for false positive reduction in mammogram mass detection", *Computers in Biology and Medicine*, vol. 57, 2015, pp. 84–95.
- [8] N. Gedik, A. Atasoy, "Performance evaluation of the wave atom algorithm to classify mammographic images", *Turk. J. Elec. Eng. & Comp. Sci.*, vol.22, 2014, pp. 957–969.
- [9] V. Chaurasia, S. Pal, "A novel approach for breast cancer detection using data mining techniques", *International Journal of Innovative Research in Computer and Communication Engineering*, vol. 2, no. 1, 2014, pp. 1-17.
- [10] L. Dora, S. Agrawal, R. Panda, A. Abraham, "Optimal breast cancer classification using Gauss–Newton representation based algorithm", *Expert Systems with Applications*, vol. 85, 2017, pp. 134-145.
- [11] N. Gedik, "Breast cancer diagnosis system via contourlet transform with sharp frequency localization and LS-SVM", *Journal of medical imaging and health informatics*, vol. 5, 2015, pp. 1–9.
- [12] W. Yang, L. Tianhui, "A Robust Feature Vector Based on Waveatom Transform for Mammographic Mass Detection," *ICVR 2018 Proceedings of the 4th International Conference on Virtual Reality*, Hong Kong, pp.133-139, 24-26 February 2018.
- [13] N. Gedik, "A new feature extraction method based on multi-resolution representations of mammograms", *Applied Soft Computing*, vol. 44, no. 1, 2016, pp. 128-133.
- [14] M.M. Jadoon, Q. Zhang, I.U. Haq, A. Jadoon, A. Basit, S. Butt, "Classification of mammograms for breast cancer detection based on curvelet transform and multi-layer perceptron", *Biomedical Research*, vol. 28, no. 10, 2017, pp. 4311-4315.

TABLE II
THE COMPARISON OF THE PROPOSED METHOD WITH PREVIOUS STUDIES

	Database	ROIs	Method	Number of features	Classification	Accuracy	Classifier
Eltoukhy et al. [16]	MIAS	322 images (207 normal, 115 abnormal, 128x128 pixel)	Curvelet	5663	N-A	95,98%	SVM
				333	B-M	97,30%	
			Wavelet	1238	N-A	95,84%	
				150	B-M	96,56%	
Gedik [13]	MIAS	228 images (114 normal, 114 abnormal, 128x128 pixel)	FFST	8776	N-A	98,29%	SVM
				66	B-M	97,39%	
	DDSM	228 images (114 normal, 114 abnormal, 128x128 pixel)	FFST	35857	N-A	96,89%	
				272	B-M	98,08%	
Eltoukhy et al. [17]	MIAS	322 images (207 normal, 115 abnormal, 128x128 pixel)	Curvelet	not given	N-A	91,19%	SVM
				not given	B-M	91,18%	
			Wavelet	not given	N-A	89,51%	
				not given	B-M	89,58%	
Proposed method	MIAS	228 images (114 normal, 114 abnormal, 128x128 pixel)	CTsFL	129	N-A	98,68%	SVM
				175	B-M	97,35%	

N-A: Normal-Abnormal, B-M: Benign-Malignant

- [15] Y. Chen, Y. Zhang, H.M. Lu, X.Q. Chen, J.W. Li, S.H. Wang, "Wavelet energy entropy and linear regression classifier for detecting abnormal breasts", *Multimed Tools Appl.*, vol. 77, 2018, pp. 3813–3832.
- [16] M.M. Eltoukhy, I. Faye, B.B. Samir, "A statistical based feature extraction method for breast cancer diagnosis in digital mammogram using multiresolution representation", *Computers in biology and medicine*, vol. 42, no. 1, 2012, pp. 123–128.
- [17] M.M. Eltoukhy, I. Faye, "An optimized feature selection method for breast cancer diagnosis in digital mammogram using multiresolution representation", *Applied Mathematics and Information Sciences*, vol. 8, no. 6, 2014, pp. 2921-2928.
- [18] D. Sehrawat, A. Sehrawat, D. Jaiswal, A. Sen, "Detection and classification of tumor in mammograms using discrete wavelet transform and support vector machine", *International Research Journal of Engineering and Technology (IRJET)*, vol. 4, no. 5, 2017, pp. 1328-1334.
- [19] Y. Lu, M.N. Do, "A new contourlet transform with sharp frequency localization", *IEEE 2006 International Conference on Image Processing*, Atlanta, Georgia, U.S.A., pp.1629-1632, 8-11 October 2006.
- [20] V. Vapnik, "The Nature of Statistical Learning Theory", Springer-Verlag New York, 1995.
- [21] V. Vapnik, "Statistical Learning Theory", New York: Wiley, 1998.
- [22] V. Vapnik, E. Levin, Y.L. Cun, "Measuring the VC-dimension of a Learning Machine", *Neural Computation*, 6(5), MIT Press, Cambridge, 1994.
- [23] H. Liu, J. Li, L. Wong, "A comparative study on feature selection and classification methods using gene expression profiles and proteomic patterns", *Genome Inf.*, vol. 13, no. 1, 2002, pp. 51–60.
- [24] <http://peipa.essex.ac.uk/info/mias.html> (20.11.2018)

BIOGRAPHIES



Nebi Gedik received his B.S. degree in Electrical and Electronics Engineering from Firat University in 2001, his PhD degrees in Electrical and Electronics Engineering from Karadeniz Technical University in 2013, and his MSc degree in 2005 from Atatürk University. He is now an Assistant Professor at University of Health Science. His research interest includes medical image and signal processing, pattern recognition and machine learning.

Improving Farm Management Information Systems with Data Mining

H. J. M. HOVING, A. KASSAHUN and C. CATAL*

Abstract— Over the past several years, farm enterprises have grown in size substantially while their number has steadily declined. As the size of their farms grow more and more farmers are deploying information systems, commonly called as Farm Management Information Systems (FMIS), to manage the day to day activities of their farms. The deployment of FMIS enable farmers to capture detailed data that can potentially be analysed by data mining tools to provide valuable information for optimizing the farm enterprises. However, data mining is generally not a common feature of many FMIS. In order to evaluate the suitability of data mining for use in FMIS, two case studies were performed using data captured in FMIS and applying various data mining algorithms. Microsoft Azure Machine Learning Studio is chosen because it provides a simple drag-and-drop visual interface that can be used by farm domain experts. In this study, two common problems were addressed in dairy farming: calving prediction of dairy cows and prediction of lactation value of milking cows. In both cases data mining models were built and experiments were run and results in both cases indicate that the required data is available from FMIS and data mining techniques provides acceptable performance. It was also shown that farm domain experts can easily use a user-friendly and drag-and-drop data mining tools with minimal initial training. Based on the insight from the two case studies and literature study, several decision problems that can be addressed with data mining such as heat prediction and lameness prediction were identified.

Index Terms— Farm management information systems, machine learning, calving prediction, lactation prediction.

I. INTRODUCTION

THE USE of data, and thus evidence, for decision making makes decisions systematic, unbiased, more accurate, and effective [1, 3, 6, 20].

While experience-based decision making by an expert may not require much data, but it definitely requires many years of observation and practice and is, therefore, very expensive. It is

HENK JAN HOVING, is with Information Technology Group, Wageningen University, Wageningen, The Netherlands, (e-mail: henk-jan.hoving@wur.nl)

 <https://orcid.org/0000-0001-7990-4237>

AYALEW KASSAHUN, is with Information Technology Group, Wageningen University, Wageningen, The Netherlands, (e-mail: ayalew.kassahun@wur.nl)

 <https://orcid.org/0000-0003-1066-7127>

CAGATAY CATAL, is with Information Technology Group, Wageningen University, Wageningen, The Netherlands, (e-mail: cagatay.catal@wur.nl)

 <https://orcid.org/0000-0003-0959-2930>

Manuscript received April 18, 2019; accepted July 19, 2019.

DOI: [10.17694/bajece.555680](https://doi.org/10.17694/bajece.555680)

also very time-consuming and labour-intensive as the expert has often to be physically present and make various measurements and observations.

When appropriate data is available, it is easier, cheaper, and often more accurate, to make decisions using computers and software systems. Moreover, the decision can be made automatically, and thus quickly. Therefore, data mining can be a very useful tool. Traditionally, many farmers make decisions mainly based on their personal experiences enhanced by a limited amount of data gathered in their paper-based diary. These decisions are called experience-based or simply expert decisions.

Recently more and more farmers are using management information systems that support simple and routine decisions. The management support systems made for farm enterprises are generally called *Farm Management Information Systems (FMIS)* [32]. FMIS generally gather a large amount of data on a daily basis from several sensors used in the farm such as milking robots and tag readers. Such a large amount of data, nowadays referred to as *big data*, can enable making non-trivial decisions, such as predicting expected calving date in dairy farming [25, 17]. In order to extract the right knowledge and patterns from big data and potentially make non-trivial decisions, several technologies including data mining, machine learning, and deep learning can be applied.

Data mining is defined as discovering new knowledge from the data [23]. It addresses a set of methods which are used to find correlations, patterns, and interesting relations between different data points in large databases. The use of data mining within FMIS can help automate decision making and even help discover new relationships among unanticipated factors. However, data mining is still not a common feature of FMIS, and it is not clear to what extent data mining can be integrated with FMIS and help address farm decision making problems and which problems it can solve effectively using the data that is commonly available within FMIS.

Data-driven decision making in comparison to expert decision making requires diverse kinds of data. An example of a decision problem that can be solved with data-driven decision making is the estimation of how long a lactation cycle of a cow will last based on the milk yield per cow [13]. Most farmers are currently not aware of how automated decision making will impact their daily life and the productivity of the cows. Naturally, there are several farmers who want to understand the underlying decision-making processes, gather the required data, deploy data mining models and take full advantage of the tools. When farmers understand how data-driven decision making works, they can make substantiated

decision on the investment of new sensors and devices and thereby, improve the state-of-the-art.

There are currently easy to use platforms for applying data mining algorithms by domain experts, of which Microsoft Azure Machine Learning Studio (MAMLS, <https://studio.azureml.net/>) is the most prominent of such tools. MAMLS provides many data mining algorithms out-of-the-box and requires no prior programming expertise to build data mining models and perform experiments. Expert users can customize MAMLS the data mining models using the R programming language but as one of the aims of this research is also to evaluate to what extent farm domain experts (specialists in farm and expert users of FMIS) can utilize data mining platforms, only the out-of-the-box algorithms of MAMLS has been used.

The uncertainties related to the above-mentioned problems make it difficult for farmers to make investments in data capturing devices and data mining tools. To address these problems, the following main research question and sub questions were formulated:

- *RQ*: How can FMIS be improved with data mining?
 - *RQ1*: Which problems at farms can be solved with data mining techniques?
 - *RQ2*: How can we these problems be mapped into data mining tasks?
 - *RQ3*: How can Azure ML Studio be applied by a domain expert?
 - *RQ4*: What kind of data is gathered at the farm and is more suitable for data mining?

The first sub question was addressed through literature search and supplemented with information from interviews and discussions with farmers. These interviews and discussions allowed us to get multiple points of view on the problems. The second sub question was addressed by using a data mining process and data mining tasks retrieved through literature search. Also, the case studies were used to help address this sub research question. The third sub-question is addressed by evaluating problems and data in detail and making several experiments by a domain expert, who is the first author of this article. The last sub question is addressed by both through literature research and discussion with different farmers, particularly, to investigate if data required for applying data mining is available in their FMIS.

To answer these research questions, the livestock sector and mainly on dairy farming were focused on. Particularly, two case studies were conducted in this research. The first case study was on the use of data mining algorithms for calving prediction problem, which is evaluated as a regression problem. The second case study focused on the estimation of milk production. Also, results of two interviews with farmers in the Netherlands are presented. Several experiments were conducted not only on public datasets, but also on our own dataset collected from a family dairy farm *Veefokbedrijf Hoving* that is associated with the first author.

The main contributions of this study are shown as follows:

- It was demonstrated that data mining provides acceptable results for calving prediction and the estimation of milk production.

- It was gathered a new dataset for predicting the milk production in a dairy farm and perform several experiments on Azure ML Studio platform.
- It was presented several tasks from dairy farming domain, which can be addressed by data mining techniques and tools.
- Based on our interviews in the Netherlands, it was concluded that farmers are already collecting sufficient data which can be used for data mining experiments.

The reminder of this paper is organized as follows: Section II explains the background and related work. Section III presents the methodology and the Section IV shows the experimental results. Section V presents the discussion and Section VI shows the conclusion and future work.

II. BACKGROUND AND RELATED WORK

A. Farm Management Information Systems (FMIS)

Essentially an FMIS is a Management Information System for farm enterprises. The key features of FMIS include financial management, reporting (also related to regulatory requirements), data acquisition, and the planning and management of farm operations, resources and people [39].

As in any other business, farmers too aim to minimize their production costs and maximize their yield [35]. The deployment of a management system helps to address this expectation by gathering, processing and synthesizing information following state-of-the-art management practices. Without an FMIS, managing a farm is a time-consuming, expensive, and labour-intensive task.

There are currently various types IoT (Internet of Things) devices such as sensors, actuators and machines used in farms. These devices gather vast amounts of data that are not captured or well-managed by the conventional FMIS systems that are widely used. Unfortunately, therefore, many FMIS do not make full use of the available data within farm enterprises [24].

In dairy farming a large amount of data is collected during the milking process of cows. The data includes the amount of milk, fat percentage, protein and lactose content, amount of urea, and somatic cell score. Based on this data, it is possible to predict how much milk the cows will be able to produce during the rest of the lactation cycle and make various decisions. For instance, a decision variable used in dairy farming is a *lactation value*. A lactation value is an indication for how well a cow is compared to the herd and can be computed based on the amount of fat, protein, and lactose content of the milk. An average cow has a lactation value of 100; if a cow performs worse than the average, its lactation value is less 100, otherwise its lactation value is equal to or above 100. Such information can be used to group the herd based on their milk production performance. The required data for determining lactation value is registered by software system that is associated with the milking machine.

Yet another example of decision variable is the Estimated Breeding Value (EBV). EBV is the combination of characteristics that are measured and marked by a livestock inspector. Unlike the lactation value, EBV is labour intensive activity, the results of which are registered in an FMIS. The

characteristics registered by the inspector include height, capacity, condition, legs, claws, and udder, and describe the phenotype of the cow. The goal for a farmer is to improve the EBV and the genetic superiority of individual cows and the overall herd.

The optimal use of data can be made when the various systems used at farm are integrated to the FMIS the farmer uses [20, 6]. The deployment of the systems and their integration with the FMIS requires that farmers understand what useful information can be extracted from the data and which problems the derived information can help address reliably.

B. Data Mining

Data mining field includes several techniques which make it possible to analyse large datasets in a short time. It helps to find different patterns in the data automatically but leaves the noise out [34]. Data mining has been used in many applications like market analysis and management, risk analyses, and fraud detection [23]. During the processing of the data, a part of the data is used to “train” the algorithm. This data is called the training data, which lets the algorithm to determine its parameters in order to identify different patterns, correlations, and anomalies within the data. After the training the algorithm, it is necessary to test it using the rest of the data set in order to make sure that training is reliable. After training and testing, the algorithm can be transformed into a software module and deployed as a web service in cloud platform.

Learning algorithms in data mining are divided into the following four categories:

1. *Supervised learning*: For supervised learning, training data should be labelled by considering different classes which exist in the data.
2. *Unsupervised learning*: In unsupervised learning, there is no output value, and only features / independent variables exist. The goal of unsupervised learning is to get a quick answer about the data [15]. It aims to infer a hidden structure in the data [26]. It also omits the noise which can influence the model [34].
3. *Semi-supervised learning*: Most of the time it is difficult to use supervised learning because not all the data is labelled. It is possible to give labels to all those data points, however, in practice this will be very expensive and time-consuming [14]. Semi-supervised learning algorithms can be used if there is very limited labelled data such as 10-20% and the remaining part of the dataset (80-90%) is unlabelled. The algorithm first tries to assign pre-labels to the unlabelled data and then, after several iterations, these labels are used in conjunction with the labelled data.
4. *Reinforcement learning*: During reinforcement learning, the agent observes the environment, performs some actions within the environment, and gets some rewards (either positive or negative) due to these actions [33].

One of the well-known and widely applied process is *Cross-Industry Standard Process for Data Mining (CRISP-DM)*. The iterative process model contains several steps as shown in Figure 1 [26]. All the steps of this process are explained as follows and it starts with the business understanding step.

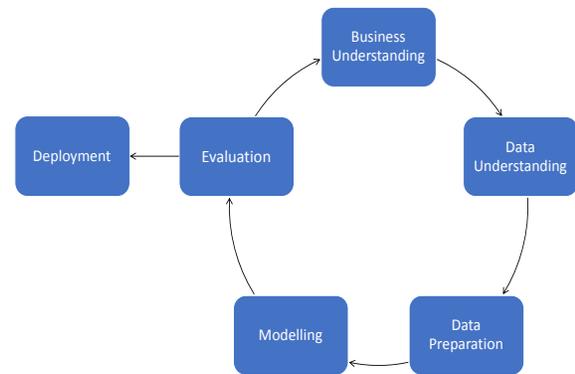


Fig. 1 CRISP-DM process (adapted from Provost & Fawcett [26])

1. *Business understanding*: This step helps to transform the business problem into one of data mining tasks.
2. *Data understanding*: After transforming the business problem into one of the data mining tasks, a dataset is required to build the models. In this step, the main task is to understand the format of the data and know how the relevant data can be collected.
3. *Data preparation*: In this step, all the raw data is prepared for modelling, data is cleaned up if necessary, and data is converted into the appropriate format.
4. *Modelling*: During this step, one of the learning algorithms is applied.
5. *Evaluation*: During the evaluation phase, the dataset is generally divided into two parts: a training set and testing set. This evaluation approach is called *hold-out*. In addition to the hold-out evaluation strategy, there is K-fold cross-validation approach which divides the dataset into K folds, which is mostly selected as 10. In each iteration, training is performed based on the (K-1) folds and the testing is done on the resting fold. This is repeated K times, and the average of values are calculated to determine the performance of the model.
6. *Deployment*: After the model is determined at the end of the evaluation phase, the model should be deployed into a web server or a cloud platform to let client applications easily access to this prediction model.

C. Data Mining Tasks

In data mining, there are hundreds of data mining algorithms. However, compared to the number of data mining algorithms, the number of data mining tasks are very less and very limited. There exist nine data mining tasks in data mining [26]. These tasks are explained as follows and are adaptations from [26]:

1. *Classification*: This task predicts for each individual data point in which class that point belongs to.
2. *Regression*: In regression, the aim is to predict a numerical a value which is actually a dependent variable.
3. *Similarity matching*: The goal of similarity matching is to find items (such as people, animals or other objects) that are similar to each other.

4. *Clustering*: Clustering techniques group individuals based on their common features, but are not driven by any specific purpose.
5. *Co-occurrence grouping*: Co-occurrence grouping finds associations between individuals, which is also known as *market-basket analysis*, *frequent item set mining*, and *association rule discovery*.
6. *Profiling*: Profiling, known as anomaly detection, characterizes the typical behaviour of an individual or a group and finds the abnormality.
7. *Link prediction*: Link prediction predicts the new links between groups or individuals.
8. *Data reduction*: In data reduction, most of the time the number of features of the dataset is reduced, but it is also possible to reduce the data points if they are detected as noisy instances.
9. *Causal modelling*: Causal modelling tries to understand what kind of events or actions influence certain events.

D. Related Work

Shahinfar et al. [29] predicted insemination outcomes in Holstein dairy cattle using data mining algorithms. Their objective was to create a user-friendly tool to help farmers make decisions. Rutten et al. [28] built a model to estimate the impact of the change of the first insemination. Fenlon et al. [12] worked on the estimation of insemination outcome by using data mining techniques for Irish dairy cows. Later, they used data mining algorithms to estimate the alignment for insemination in farms which use a seasonal calving system [12]. They stated that it is important that peak grass availability with peak lactating cow energy balances are aligned. Mahmoud et al. [22] used the RumiWatch device for estimating the calving time. By using the data about the rumination behaviour, it was possible to estimate the calving time. Zaborski et al. [38] used data mining techniques to detect dystocia in dairy cattle. Their approach gives a clear view which heifers get troubles during calving.

Borchers et al. [4] applied artificial neural networks for calving prediction. For gathering data, they used the HR Tag device of the company SCR Engineers and the data contained information about the rumination behaviour, neck activity, number of steps, and lying time.

For estimating the sickness in herds, data mining has been previously used. Different algorithms were used to predict the sick cows by using data containing grazing, standing, and rumination behaviour [27]. Yazdanbakhsh et al. [37] performed a research on sickness in an Alberta feedlot. They were capable of finding illness up to seven days in advance by using data mining techniques and an inferential sensor. Caraviello et al. [5] used several data mining methods to estimate the effect of factors on the reproduction of lactating Holstein cows. Kim and Heald [21] applied data mining techniques to get an idea of which bacterial was causing mastitis in a dairy herd. Kamphuis et al. [19] applied decision trees for the detection of clinical mastitis in farms using automatic milking. Alsaad et al. [2] worked on the lameness prediction problem, and predicted accurately the early state of lameness of cows.

Heat detection is also important in dairy businesses. Heat detection using accelerometers and unsupervised learning has been performed by Shahriar et al. [30]. Vanrell et al. [35] applied accelerometers to predict the heat events in dairy cows using decision trees algorithms. Chung et al. [7] used small microphones to collect data about the heat events and applied data mining to build models.

III. METHODOLOGY

This research required data gathering for data mining, experimenting with various data mining methods and tasks, and finally, interviewing farmers to assess the suitability of data mining in supporting their decision making. Each of these methods are explained, starting with experimental design, followed by the data collected and then, the interviews. To design the experiments, MAMLS's graphical and experiment designer were used. Figure 2 shows the experimental design built for the calving prediction case study (Case Study-I).

The experimental design depicted in Figure 2 presents the following steps. The data set *days.csv* is uploaded to MAMLS. This dataset is split into two parts (one part is used for training and another for testing) with *Split Data* visual component. Training is performed in the *Train Model*. The inputs to Train Model are the outputs of the Split Data component and a data mining algorithm (in this case the *Neural Network Regression* algorithm). The result of the training is tested using the *Score Model* component. Besides the output of Train Model, the Score Model component requires the second output of Split Data. Finally, in *Evaluate Model* component uses the output of the Score Model to determine and present several evaluation parameters.

For regression tasks the following algorithms were applied:

- *Bayesian networks*: This algorithm applies probability theory for its calculations, which is based on random variables and their conditional independencies.
- *Decision tree*: This is one of the most simple but useful algorithms in data mining. It starts with a node at the base and extends to several leaf nodes which show the classes the tree can classify. Overfitting is one of the drawbacks of this algorithm. There are different implementations of this algorithm and one of them is boosted decision trees. Another example is the Random Forest algorithm which uses multiple trees for the decision.
- *Linear regression*: This algorithm aims to find a relationship between inputs and the output and it can deal with multiple input variables.
- *Neural Networks (NN)*: NN is an algorithm inspired by the biological neural networks which are complex systems that contain many neurons. There are different topological models of NNs such as multi-layer perceptron and recurrent neural networks. Nowadays, deep learning algorithms, which are based on NNs, are widely used in different application domains.
- *Poisson Regression*: The Poisson regression, which is a type of generalized linear model, is based on the Poisson distribution.

The suitability of a data mining algorithm for a given problem and dataset is determined by evaluation parameters. The

following evaluation parameters have been used for classification and regression problems in this study:

- **Mean absolute error:** It measures the error between two continuous variables such as X and Y. It can take values between zero and infinity.
- **Root mean squared error:** This value provides the standard deviation of the residuals. It is between zero and infinity.
- **Relative squared error:** It gives a normalized squared error which means that the relative squared error takes the total squared error and normalizes it by dividing with the total squared error of the prediction. The result can be between zero and infinity.
- **Relative absolute error:** This is very similar to the relative squared error, but it is not squared. To make the outcomes positive, the formula contains absolute sign, therefore the results are between zero and infinity.
- **Coefficient of determination:** This parameter is also known as R^2 . Adding more variables to the model will not necessarily reduce the coefficient of determination, but could indeed reduce prediction accuracy by introducing prediction variance [18]. It is between 0 and 1 and when it is near to 1, the model is said to be perfect. If it is negative, it means that the model cannot represent the underlying data.
- **Overall accuracy:** This parameter provides the non-weighted accuracy of all the classes combined.
- **Average accuracy:** The average accuracy is the weighed accuracy for all the classes, which indicates the goodness of a classification model as a proportion of true results to total cases.
- **Precision:** Precision provides an indication for the correct positive observations. This can be calculated from a confusion matrix, the following formula is used:

$$\text{Precision} = \frac{\text{true positives}}{\text{true positives} + \text{false positives}}$$
- **Recall:** The recall is a synonym for sensitivity, which is the ratio of correctly predicted positive events. The formula is shown as follows:

$$\text{Recall} = \frac{\text{true positive}}{\text{true positives} + \text{false negatives}}$$

A. Experimental Design for Calving Prediction

As shown in Figure 2, an experiment has been prepared for the calving prediction problem. In this experiment, 80% of the dataset has been used for the training and the rest has been used for testing. In Figure 3, it is shown how we can design an experiment to perform 10-fold cross validation analysis. This time *Partition and Sample* component was used instead of *Split Data* component.

B. Experimental Design for Predicting Lactation Value

For the second case study, data obtained from dairy farm Veefokbedrijf Hoving was used. At the beginning, it was aimed to estimate the lactation value of individual cows using regression algorithms. However, lactation values obtained were not sufficiently accurate; therefore, it was resorted to building a multi-class classification model.

The data points were divided into three classes: bad-0 (lactation values below 95), neutral-1 (lactation values between 95 and 105), and good-2 (lactation values over 105). In this experiment, data was obtained as a PDF file, which is

the only data format some FMIS use for data exchange. Therefore, first the data had to be transformed into a csv file.

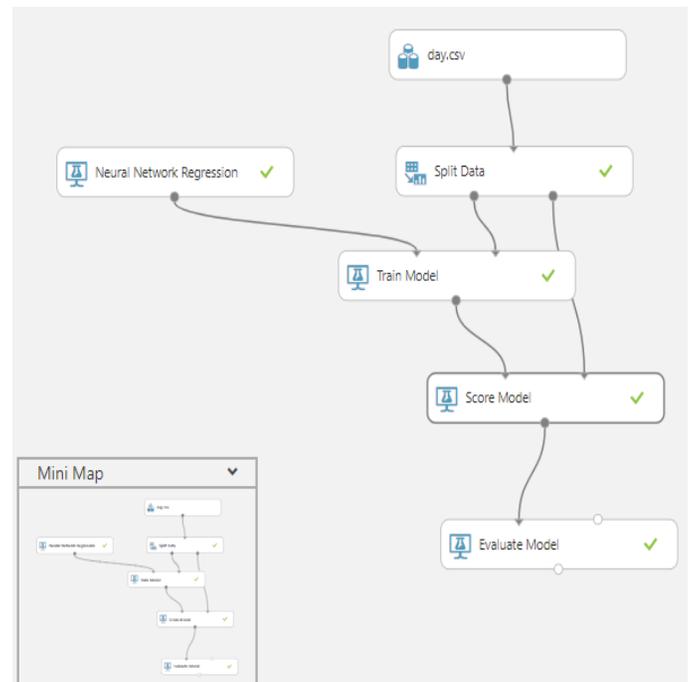


Fig. 2. Experiment for calving prediction case study in Azure ML Studio

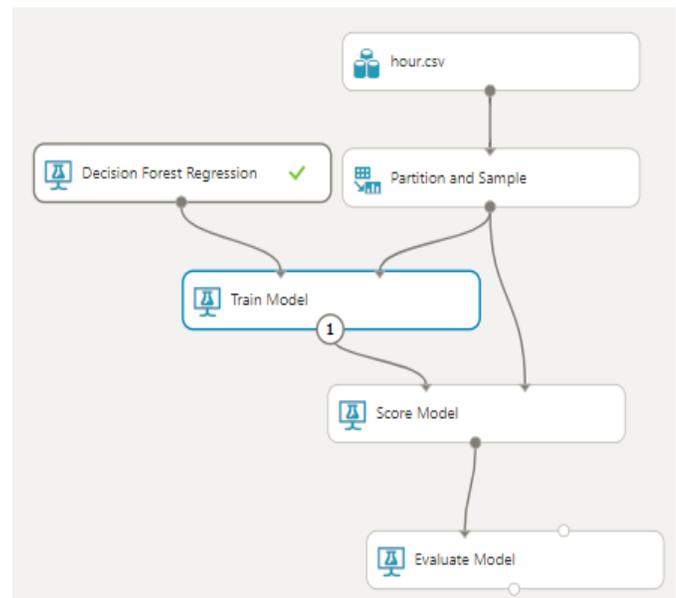


Fig. 3. Experiment for 10-fold cross-validation

Since many features were not useful, *Select Columns in Dataset* component has been used during the experimental design as shown in Figure 4. The following useful columns were selected: *Milk yield, % fat, % protein, % lactose, urea, days after calving, age, lactation number, and class label*. The dataset was split into 80% training data and 20% testing data. First multiclass neural network algorithm has been used for the analysis, and then, multiclass forest, multiclass decision jungle, and multiclass logistics regression algorithms were applied.

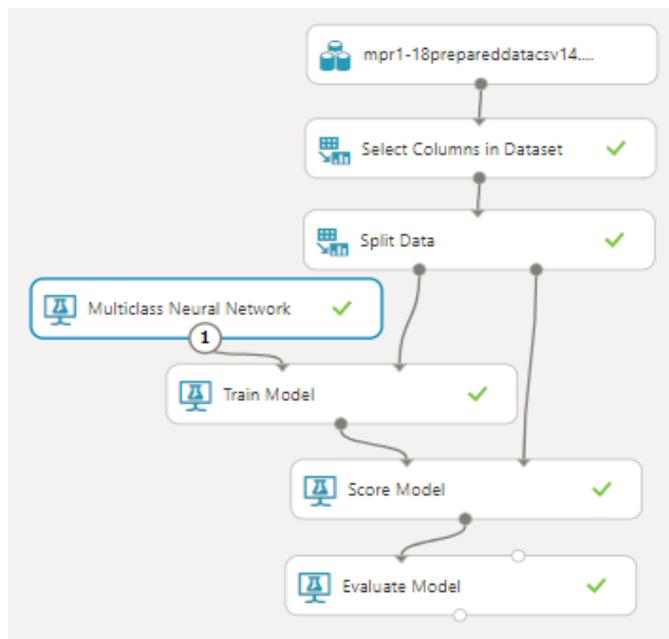


Fig. 4. Experiment for prediction lactation value (multi-class classification)

To get more information about the quality of lactation values, authors contacted CRV (<https://www.crv4all.nl/>), the company that managed the data. The company explained that the lactation values are partly measured and partly estimated. The company used basic lactation curve and added measurement values from individual cows. These measurements contain data about milk yield, percentage fat, percentage protein, percentage lactose, and the amount of urea. Furthermore, information about the cow, such as age, calving date, and lactation number are used with the model the company used to estimate lactation value. For estimating productions between two measurements, interpolation is applied [11]. By using the model output of the CRV, in this case study authors were measuring how well data mining results are matching with the classical methods of computing the productivity of a cow in the future and thereby showing the potential of data mining.

After building a classification model, it was aimed to build a regression model for this dataset. During the regression experiments, linear regression, Bayesian linear regression, boosted decision tree regression, decision forest regression, Poisson regression, and neural network algorithm were used.

C. Data Collection

For calving prediction experiments, two datasets (day.csv and hour.csv) shared by Borchers et al. [4] were used. One of them contains data for the prediction on which day the cow will calve and the other one contains data for estimating on which hour of the day the cow will calve. The data contains minute-by-minute data log of activity, lying, and ruminating behaviour of every dairy cattle. This data is gathered using sensors that are capable of measuring all these different activities. The data contains the monthly values milk production of all the cows, including the fat, protein, lactose, protein, and urea contents of the milk. Dataset can be accessed from the following public github link: [https://github.com/Mrborchers/Machine-learning-](https://github.com/Mrborchers/Machine-learning-based-calving-prediction-from-activity-lying-and-rumination-behaviors)

[based-calving-prediction-from-activity-lying-and-rumination-behaviors](https://github.com/Mrborchers/Machine-learning-based-calving-prediction-from-activity-lying-and-rumination-behaviors)

D. Interviews

The dairy farm Vee fokbedrijf Hoving has a herd of 60 milking cows and an old-school milking parlour. The interviewee measures the milk production with a monthly MPR (milk production registration). At the time of the interview, the farmer was content with the FMIS he has in place but would also be happy if he has a way of predicting certain events because predicting events would make his life easier. Therefore, in order to understand the value of prediction, it was asked to him how much money he would spend per cow if such a system can be built for the farm. He replied that he would spend €500 per cow. Several tasks were listed for him and he provided the following order of priority: prediction of sickness, heat detection, calving prediction, and lameness detection.

The second interviewee works part-time in a family farm which has a herd of 50 cows and a milking robot. Therefore, he has daily information about the milk production and concentrate consumption of every individual cow. The data about milk production is comparable with data from MPR, however it is daily instead of monthly. Furthermore, the cows have activity sensors on their neck and he plots crowing curves for the calves. At the time of the interview, he uses data from milking robot [9, 10] to make decisions about the heat activities, sickness, and lameness. Using the calve growing curves, he follows the growth of all the calves and if they do not grow optimally he changes the feeding patterns. He would like to see some reports that show which cows will have a heat. Though visual heat detection is manageable for 50 cows, it will be difficult when the number of cows reaches to 200. In the current circumstances, he would like to identify cows in heat that needs special attention—*i.e.* he wants to identify the outliers in heat behaviour. When it was asked how much he can spend for each cow, he responded that this is dependent on how much the sensors would return and how the system will present the necessary early events and warnings. His ranking for the tasks was as follows: prediction of sickness, heat detection, lameness, and calving prediction. Based on our interviews, it was observed that the prediction of sickness and heat detection have higher priority compared to the other tasks.

IV. EXPERIMENTAL RESULTS

A. Case Study 1- Calving Prediction

In Table 1, the values regarding the five evaluation parameters for the 6 regression algorithms are presented. Of the 6 different regression algorithms linear regression and Bayesian linear regression algorithms resulted a perfect coefficient of determination (which is 1), which means that the two algorithms work perfectly on the given datasets. Also, the mean absolute errors for these two algorithms are very low. The other regression algorithms are not as good as the first two as can be seen from the values of the coefficient of determination (less than one) and the mean absolute errors (greater than 1). In addition to the hold-out approach, Cross-

Validation (CV) analysis was also used. The results of cross validation are presented in Table 2. A similar performance for linear regression and Bayesian linear regression algorithms are observed, both of which reach to the value 1 for coefficient of determination parameter.

The results for the prediction in which hour the cows will calf are shown in Table 3. Based on this table, it can be stated that compared to the other algorithms the Boosted Decision Tree Regression and Decision Forest Regression algorithms perform the best. In Table 4 the results of predicting the hour of calving using 10-fold CV are presented. The results are very similar to the results of Table 3. Boosted Decision Tree Regression and Decision Forest Regression algorithms perform better than the other algorithms.

TABLE I.

CALVING PREDICTION RESULTS BASED ON DAILY DATA						
Algorithms →	LR	BayR	BooDTR	DFR	NN	PR
MAbsE	0.000247	0.000026	1.43	2.2	2.51	2.91
RMSqE	0.00032	0.000034	1.89	2.86	3.32	3.42
RelAbsE	0.000065	0.000007	0.375	0.549	0.658	0.765
RelSqE	0.	0.	0.184	0.392	0.566	0.619
CoefDet	1.	1.	0.816	0.608	0.434	0.381

LEGEND:

MAbsE	Mean Absolute Error
RMSqE	Root Mean Squared Error
RELABS E	Relative Absolute Error
RELSqE	Relative Squared Error
COEFD E T	Coefficient of determination
LR	Linear regression
BAYLR	Bayesian Linear Regression
BOODTR	Boosted Decision Tree Regression
DFR	Decision Forest Regression
NN	Neural Network
PR	Poisson Regression

TABLE II.

CALVING PREDICTION BASED ON DAILY DATA USING 10-FOLD CV						
Algorithms →	LR	BayR	BooDTR	DFR	NN	PR
MAbsE	0.000288	0.000196	1.56	2.23	2.7	3.79
RMSqE	0.000456	0.000346	2.1	2.83	3.53	4.16
RelAbsE	0.000076	0.000052	0.412	0.591	0.711	0.904
RelSqE	0.	0.	0.231	0.422	0.654	0.817
CoefDet	1.	1.	0.769	0.578	0.346	0.183

TABLE III.

CALVING PREDICTION RESULTS BASED ON HOURLY DATA						
Algorithms →	LR	BayR	BooDTR	DFR	NN	PR
MAbsE	0.263	0.202	0.	0.0499	0.23	1.47
RMSqE	0.254	0.249	0.	0.11	0.281	1.78
RelAbsE	0.295	0.288	0.	0.0701	0.329	0.24
RelSqE	0.0921	0.0891	0.	0.0172	0.113	0.0648
CoefDet	0.908	0.911	1.	0.983	0.887	0.935

TABLE IV.

CALVING PREDICTION BASED ON HOURLY DATA - 10-FOLD CV						
Algorithms →	LR	BayR	BooDTR	DFR	NN	PR
MAbsE	0.217	0.214	0.	0.0421	0.243	1.6
RMSqE	0.265	0.262	0.	0.094	0.298	1.93
RelAbsE	0.325	0.321	0.	0.063	0.363	0.284
RelSqE	0.106	0.104	0.	0.0134	0.134	0.0851
CoefDet	0.894	0.896	1.	0.987	0.866	0.915

B. Case Study 2- Predicting Lactation Value

For the second case study, experiments using multi-class classification algorithms and then based on regression algorithms were performed. For the classification models, the overall accuracy parameter can be used for selecting the right model. For the regression models, the coefficient of determination parameter is the indicator for selecting the right model. In Tables 5, the results of classification algorithms based on single measurement are presented and in Table 6, the results of regression algorithms based on single measurement are shown. In Table 7 and 8, the results of regression algorithms based on multiple measurements and classification results based on multiple measurements are shown, respectively. Based on these results, it can be stated that the performance of the multi-class classification models improved when multiple measurements are added to the dataset. As shown in Table 8, the average accuracy of the multiclass decision tree is around 0.88 which is quite high. This indicates that it can be determined which cows perform relatively poor, and which ones perform good in terms of milk production. The performance of regression algorithms, on the other hand, is not high enough for practical use, therefore, it is recommended the use of classification models for decision making at dairy farms. However, note that since the data was generated based on the lactation value computed by CRV (instead of values that are measured), sufficient insights to what extent the dependent variable is accurate are not known, and as a result it is not expected to reach a high coefficient of determination values for regression models.

Table 5 shows the results of the lactation value estimation with a classification model using only single measurement. The performance is not acceptable due to the low overall accuracy value. The best models are multiclass decision forest and multiclass neural network having the same overall accuracy value (0.625).

TABLE V.

PREDICTING LACTATION VALUE (CLASSIFICATION RESULTS)

Used features Algorithms →	Milk yield, % fat, % protein, % lactose, urea, days after calving, age, lactation number, classification			
	McDF	McDJ	McLR	McNN
OA	0.625	0.583	0.542	0.625
AVA	0.75	0.722	0.694	0.75
MICAVP	0.625	0.583	0.542	0.625
MACAVP	0.622	0.575	0.432	NaN
MICAVRC	0.625	0.583	0.542	0.625
MACAVRC	0.625	0.569	0.515	0.524

LEGEND:

OA	OVERALL ACCURACY
AVA	AVERAGE ACCURACY
MICAVP	MICRO-AVERAGED PRECISION
MACAVP	MACRO-AVERAGED PRECISION
MICAVRC	MICRO-AVERAGED RECALL
MACAVRC	MACRO-AVERAGED RECALL
McDF	Multiclass Decision Forest
McDJ	Multiclass Decision Jungle
McLR	Multiclass Logistic Regression
McNN	Multiclass Neural Network

TABLE VI.

PREDICTING LACTATION VALUE (REGRESSION RESULTS)

Used features Algorithms →	Milk yield, % fat, % protein, % lactose, urea, days after calving, age, lactation number, lactation value					
	LR	BayR	BooDTR	DFR	NN	PR
MABSE	11.8	15.1	13.1	12.1	12.3	16.6
RMSqE	20.2	19.9	18.2	20.3	18.5	29.9
RelAbsE	0.643	0.826	0.714	0.66	0.674	0.908
RelSqE	0.475	0.461	0.386	0.48	0.398	1.04
CoefDet	0.525	0.539	0.614	0.52	0.602	0.0398

The accuracies of lactation value estimation with regression algorithms are shown in Table 6. The best model is boosted decision tree regression with a coefficient of determination of 0.61. The second best performing one is neural network algorithm. The low values of the coefficient of determination indicate that the results are generally unacceptable.

In Table 7, the results of estimation accuracies of the lactation value with regression algorithms are presented, but this time utilizing multiple measurements. As it can be seen, the algorithms performed poorly. This indicates that these models cannot predict the lactation value accurately. For boosted decision tree regression and decision forest regression the values of coefficient of determination are positive, but still very low. Compared to the Table 6, adding more data has not improved the model performance in this case.

The results for estimating the lactation value based on classification models are presented in Table 8. In this case, multiple measurements have been used. The overall accuracy increased when more measurements are added. The best performing model is multiclass decision forest in this case.

TABLE VII.

RESULTS WITH MULTIPLE MEASUREMENTS (REGRESSION)

Used features Algorithms →	Milk yield, % fat, % protein, % lactose, urea, age, average lactation value					
	LR	BayR	BooDTR	DFR	NN	PR
MABSE	12.8	11.8	8.05	8.33	10.6	12.2
RMSqE	17.3	15.8	10.3	10.8	13.7	13.8
RelAbsE	1.16	1.06	0.729	0.755	0.96	1.1
RelSqE	1.75	1.46	0.622	0.676	1.09	1.12
CoefDet	-0.751	-0.455	0.378	0.324	-0.0894	-0.117

TABLE VIII.

RESULTS WITH MULTIPLE MEASUREMENTS (CLASSIFICATION)

Used features Algorithms →	Milk yield, % fat, % protein, % lactose, urea, age, average classification			
	McDF	McDJ	McLR	McNN
OA	0.818	0.682	0.727	0.636
AVA	0.879	0.788	0.818	0.758
MICAVP	0.818	0.682	0.727	0.636
MACAVP	0.73	0.583	0.678	0.571
MICAVRC	0.818	0.682	0.727	0.636
MACAVRC	0.857	0.762	0.81	0.952

V. DISCUSSION

In this study it was set out which problems at farm data mining can address and if farm domain experts can get the required data from an FMIS and translate them into data mining problems. It was also shown that data mining tools can be utilized by farm domain experts with little training. However, due to the empirical nature of this study, potential threats to validity must be addressed [31]. The four validity dimensions introduced by [36] are addressed, which in turn follow the work of Cook et al. [8]. The four validity dimensions we address are *conclusion*, *internal*, *external*, and *construct* validity.

To address conclusion validity threat, several evaluation parameters were used in this study. By presenting all the major evaluation parameters it is considered that the risk of other researchers applying a different evaluation parameter that will lead to a different conclusion about performance is reduced. In addition to the dataset from an FMIS for predicting the milk production, two public datasets for calving prediction were used. To strengthen the conclusion, additional experiments using more datasets from diverse FMIS are needed. Regarding the construct validity, the features in the public datasets were used as-is because those features have previously been used by the other researchers successfully. With regard to the dataset from an FMIS, the features were selected based on the domain knowledge. In order to minimize the possibility that different researchers reaching different performance results, various combination of features were investigated. External validity threat is about the limit of generalizability of the results of a case study [16]. Since the datasets are limited—one dataset belongs to only one farm in the Netherlands and the two public datasets belong only to one country, the performance of the models might change when

more datasets become available from different countries. To limit internal validity threats, a number of algorithms were applied for each problem and therefore, results are not limited by one or a few algorithms. The experiments are limited with the algorithms which exist in MAMLS and other researchers might achieve better performance with new algorithms that do not present in MAMLS.

VI. CONCLUSION

It was shown that a number of decisions at farm enterprises that until today required human expert can be supported or automated by data mining algorithms using data available in today's FMIS. For instance, high performance was achieved for the calving prediction problem. To make a hard conclusion on predicting the lactation value, more accurate lactation values are required. The other identified problems in this study are heat detection and lameness prediction. It was demonstrated that MAMLS can be used by a farm domain expert easily. It is planned to let a number of domain experts apply MAMLS on their own datasets in future research to reach more conclusive results.

(1) References

- [1] Albersmeier, F., Schulze, H., Jahn, G., & Spiller, A. (2009). The reliability of third-party certification in the food chain: From checklists to risk-oriented auditing. *Food Control*, 20(10), 927-935. doi:https://doi.org/10.1016/j.foodcont.2009.01.010
- [2] Alsaad, M., Römer, C., Kleinmanns, J., Hendriksen, K., Rose-Meierhöfer, S., Plümer, L., & Büscher, W. (2012). Electronic detection of lameness in dairy cows through measuring pedometric activity and lying behavior. *Applied Animal Behaviour Science*, 142(3), 134-141. doi:10.1016/j.applanim.2012.10.001
- [3] Azzaro, G., Caccamo, M., Ferguson, J. D., Battiato, S., Farinella, G. M., Guarnera, G. C., . . . Licitra, G. (2011). Objective estimation of body condition score by modeling cow body shape from digital images. *J Dairy Sci*, 94(4), 2126-2137. doi:10.3168/jds.2010-3467
- [4] Borchers, M. R., Chang, Y. M., Proudfoot, K. L., Wadsworth, B. A., Stone, A. E., & Bewley, J. M. (2017). Machine-learning-based calving prediction from activity, lying, and ruminating behaviors in dairy cattle. *J Dairy Sci*, 100(7), 5664-5674. doi:10.3168/jds.2016-11526
- [5] Caraviello, D. Z., Weigel, K. A., Craven, M., Gianola, D., Cook, N. B., Nordlund, K. V., . . . Wiltbank, M. C. (2006). Analysis of Reproductive Performance of Lactating Cows on Large Dairy Farms Using Machine Learning Algorithms. *J Dairy Sci*, 89(12), 4703-4722. doi:10.3168/jds.S0022-0302(06)72521-8
- [6] Choubey, M. K. (2011). IT Infrastructure and Management (For the GBTU and MMTU): Pearson Education India.
- [7] Chung, Y., Lee, J., Oh, S., Park, D., Chang, H., & Kim, S. (2013). Automatic detection of cow's oestrus in audio surveillance system. *Asian-Australasian journal of animal sciences*, 26(7), 1030.
- [8] Cook, T. D., Campbell, D. T., & Day, A. (1979). *Quasi-experimentation: Design & analysis issues for field settings* (Vol. 351): Houghton Mifflin Boston.
- [9] Delaval. (14-3-2018a). BCS
- [10] Delaval. (14-3-2018b). View BCS.
- [11] Delta, C. (February 1998). *lactatieproductie en 305 dagenproductie. Handboek NRS*.
- [12] Fenlon, C., O'Grady, L., Doherty, M., Butler, S., Shalloo, L., & Dunning, J. (2016, 12-15 Dec. 2016). Regression Techniques for Modelling Conception in Seasonally Calving Dairy Cows. Paper presented at the 2016 IEEE 16th International Conference on Data Mining Workshops (ICDMW).
- [13] Grzesiak, W., Błaszczyk, P., & Lacroix, R. (2006). Methods of predicting milk yield in dairy cows—Predictive capabilities of Wood's lactation curve and artificial neural networks (ANNs). *Computers and Electronics in Agriculture*, 54(2), 69-83.
- [14] Hady, M. F. A., & Schwenker, F. (2013). Semi-supervised Learning. In M. Bianchini, M. Maggini, & L. C. Jain (Eds.), *Handbook on Neural Information Processing* (pp. 215-239). Berlin, Heidelberg: Springer Berlin Heidelberg.
- [15] Hastie, T., Tibshirani, R., & Friedman, J. (2009). *Unsupervised learning The elements of statistical learning* (pp. 485-585): Springer.
- [16] Hosseini, S., Turhan, B., & Mäntylä, M. (2017). A benchmark study on the effectiveness of search-based data selection and feature selection for cross project defect prediction. *Information and Software Technology*.
- [17] Husemann, C., & Novković, N. (2014). *FARM MANAGEMENT INFORMATION SYSTEMS: A CASE STUDY ON A GERMAN MULTIFUNCTIONAL FARM*.
- [18] James, G., Witten, D., Hastie, T., & Tibshirani, R. (2017). *An Introduction to Statistical Learning with Applications in R*. doi:DOI 10.1007/978-1-4614-7138-7
- [19] Kamphuis, C., Mollenhorst, H., Feelders, A., Pietersma, D., & Hogeveen, H. (2010). Decision-tree induction to detect clinical mastitis with automatic milking. *Computers and Electronics in Agriculture*, 70(1), 60-68. doi:https://doi.org/10.1016/j.compag.2009.08.012
- [20] Kim, C.-H., Weston, R. H., Hodgson, A., & Lee, K.-H. (2003). The complementary use of IDEF and UML modelling approaches. *Computers in Industry*, 50(1), 35-56.
- [21] Kim, T., & Heald, C. W. (1999). Inducing inference rules for the classification of bovine mastitis. *Computers and Electronics in Agriculture*, 23(1), 27-42. doi:https://doi.org/10.1016/S0168-1699(99)00003-4
- [22] Mahmoud, F., Christopher, B., Maher, A., Jürg, H., Alexander, S., Adrian, S., & Gaby, H. (2017). Prediction of calving time in dairy cattle. *Animal Reproduction Science*, 187, 37-46. doi:https://doi.org/10.1016/j.anireprosci.2017.10.003
- [23] Han, J., Pei, J., & Kamber, M. (2011). *Data mining: concepts and techniques*. Elsevier.
- [24] Nikkilä, R., Seilonen, I., & Koskinen, K. (2010). Software architecture for farm management information systems in precision agriculture. *Computers and Electronics in Agriculture*, 70(2), 328-336. doi:https://doi.org/10.1016/j.compag.2009.08.013
- [25] Paraforos, D. S., Vassiliadis, V., Kortenbruck, D., Stamkopoulos, K., Ziogas, V., Sapounas, A. A., & Griepentrog, H. W. (2017). Multi-level automation of farm management information systems. *Computers and Electronics in Agriculture*, 142, 504-514. doi:https://doi.org/10.1016/j.compag.2017.11.022
- [26] Provost, F., & Fawcett, T. (2013). *Data Science for Business: What you need to know about data mining and data-analytic thinking*: " O'Reilly Media, Inc."
- [27] Rahman, A., Smith, D. V., Little, B., Ingham, A. B., Greenwood, P. L., & Bishop-Hurley, G. J. (2018). Cattle behaviour classification from collar, halter, and ear tag sensors. *Information Processing in Agriculture*, 5(1), 124-133. doi:https://doi.org/10.1016/j.inpa.2017.10.001
- [28] Rutten, C. J., Steeneveld, W., Vermooij, J. C. M., Huijps, K., Nielen, M., & Hogeveen, H. (2016). A prognostic model to predict the success of artificial insemination in dairy cows based on readily available data. *J Dairy Sci*, 99(8), 6764-6779. doi:10.3168/jds.2016-10935
- [29] Shahinfar, S., Page, D., Guenther, J., Cabrera, V., Fricke, P., & Weigel, K. (2014). Prediction of insemination outcomes in Holstein dairy cattle using alternative machine learning algorithms. *J Dairy Sci*, 97(2), 731-742. doi:10.3168/jds.2013-6693
- [30] Shahriar, M. S., Smith, D., Rahman, A., Freeman, M., Hills, J., Rawnsley, R., . . . Bishop-Hurley, G. (2016). Detecting heat events in dairy cows using accelerometers and unsupervised learning. *Computers and Electronics in Agriculture*, 128, 20-26. doi:https://doi.org/10.1016/j.compag.2016.08.009
- [31] Šmite, D., Wohlin, C., Gorschek, T., & Feldt, R. (2010). Empirical evidence in global software engineering: a systematic review. *Empirical software engineering*, 15(1), 91-118.
- [32] Sørensen, C. G., Pesonen, L., Bochtis, D. D., Vougioukas, S. G., & Suomi, P. (2011). Functional requirements for a future farm management information system. *Computers and Electronics in Agriculture*, 76(2), 266-276. doi:https://doi.org/10.1016/j.compag.2011.02.005

- [33] Sutton, R. S., & Barto, A. G. (1998). Reinforcement learning: An introduction (Vol. 1): MIT press Cambridge.
- [34] Valletta, J. J., Torney, C., Kings, M., Thornton, A., & Madden, J. (2017). Applications of machine learning in animal behaviour studies. *Animal Behaviour*, 124, 203-220.
- [35] Vanrell, S. R., Chelotti, J. O., Galli, J., Rufiner, H. L., & Milone, D. H. (2014). 3d acceleration for heat detection in dairy cows. Paper presented at the XLIII Jornadas Argentinas de Informática e Investigación Operativa (43JAIO)-VI Congreso Argentino de AgroInformática (CAI)(Buenos Aires, 2014).
- [36] Wohlin, C., Runeson, P., Host, M., Ohlsson, M., Regnell, B., & Wesslen, A. (2000). Experimentation in software engineering: an introduction. 2000: Kluwer Academic Publishers.
- [37] Yazdanbakhsh, O., Zhou, Y., & Dick, S. (2017). An intelligent system for livestock disease surveillance. *Information Sciences*, 378, 26-47. doi:https://doi.org/10.1016/j.ins.2016.10.026
- [38] Zaborski, D., Proskura, W. S., Grzesiak, W., Szatkowska, I., & Jędrzejczak-Silicka, M. (2017). Use of random forest for dystocia detection in dairy cattle. *Applied Agricultural and Forestry Research*, 147.
- [39] Tummers, J., Kassahun, A., & Tekinerdogan, B. (2019). Obstacles and features of Farm Management Information Systems: A systematic literature review. *Computers and Electronics in Agriculture*, 157, 189-204.

Appendix- Interview Questions

- What do you measure in your farm?
- What kind of features do you combine to make decisions?
- Do you try to make more decisions by yourself based on the measured data in your farm?
- What kind of data is missing in your farm?
- How much money would you invest per cow for a new system with several features?

Rank the following tasks:

- List from the most important to less: heat detection, lameness, calving prediction, sickness prediction

transparency and traceability systems, and applying data science in the agriculture and food supply chains. He has been involved in a number of EU sponsored research projects and supervises a large number of students on diverse thesis research topics.

CAGATAY CATAL is a member of Information Technology



Group in Wageningen University in the Netherlands. He received his B.S. and M.S. degrees in computer engineering from the Istanbul Technical University, Istanbul, and the Ph.D. degree in computer engineering from Yildiz Technical University. He worked 8 years at the Scientific and Technological Research Council of Turkey as Senior Researcher. Later, he

worked 6 years in Istanbul Kultur University, Department of Computer Engineering as Associate Professor and Head of the Department. On January 2018, he joined the Information Technology Group of Wageningen University in the Netherlands. His research interests are data science, machine learning, software quality and testing, and experimental software engineering.

BIOGRAPHIES



HENK JAN HOVING is a master student at the Wageningen University in the Netherlands. He finished his Bachelor degree in Wageningen University and completed his thesis on the application of machine learning in life sciences domain. His research interests are big data analytics, data mining, farm management

information systems, data science, and life sciences.



AYALEW KASSAHUN has a PhD in knowledge systems from Wageningen University in the Netherlands where he currently works as a researcher and lecturer. Before joining the university, he worked for five years as a software engineer, first for Baan company (now Infor), a large ERP developer, and later

for InfoRay, a developer of a business intelligence software for the banking and insurance sector. His research publications and interests are diverse, including software architecture, management information systems, business process modelling,

Model Predictive Control of an Indirect Matrix Converter with Active Damping Capability

M. GOKDAG and O. GULBUDAK

Abstract—In this paper, a model predictive control (MPC) scheme is proposed to control indirect matrix converter (IMC), which is used for three phase-to-three phase direct power conversion without any intermediate energy storage component between them. The aim in the control of current source rectifier (CSR) stage of IMC is generally to have unity power factor with relatively low total harmonic distortion (THD). The aim in the control of voltage source inverter (VSI) stage is to be able to synthesize sinusoidal load currents with desired peak value and frequency. Imposed source current MPC technique in abc frame is used for the rectifier stage. An active damping technique without requiring the selection of an appropriate value for fictitious damping resistor is also included in the proposed control scheme in order to mitigate the resonance phenomenon of lightly damped input LC filter to suppress the higher order harmonics in supply currents. Load currents with desired peak and frequency are also obtained by imposing sinusoidal currents in abc frame. Two different cost functions are combined into a single cost function without any weighting factor since both error terms are in the same nature. The switching state that minimizes this pre-defined cost function among the 24-valid switching combinations of IMC is selected and applied to converter. The proposed model predictive control with active damping method shows a good performance in terms of THD levels in supply currents even at low current demands from supply side. The proposed strategy guarantees unity power factor operation and draws sinusoidal load currents at desired peak and frequency.

Index Terms—Indirect matrix converter, ac-ac power conversion, model predictive control, active damping.

I. INTRODUCTION

MATRIX CONVERTER (MC) was introduced as an alternative to back-to-back (AC-DC-AC) converter which requires high energy storage components in the DC link prone to failure. The structure of the matrix converter which does not require any electrolytic capacitors allows it to operate with high reliability and high-power densities. The main advantages of the matrix converter can be listed as follows:

MUSTAFA GOKDAG, is with Department of Electrical and Electronics Engineering University of Karabuk University, Karabuk, Turkey, (e-mail: mgokdag@karabuk.edu.tr).

 <https://orcid.org/0000-0001-5589-2278>

OZANGULBUDAK, is with Department of Electrical and Electronics Engineering University of Karabuk, Karabuk, Turkey, (e-mail: ozangulbudak@karabuk.edu.tr).

 <https://orcid.org/0000-0001-9517-3630>

Manuscript received September 3, 2019; accepted Nov 15, 2019.

DOI: [10.17694/bajece.614528](https://doi.org/10.17694/bajece.614528)

MC inherently has bidirectional power flow, it is capable to simultaneously transform amplitude and frequency of the three-phase voltage system, sinusoidal input current and output voltage can be synthesized with suitable modulation techniques such as space vector modulation, and input power factor can be independently controlled. Because of these advantages, a great interest from academia and industry focused onto the application of matrix converters in order to realize highly compact three-phase AC drives for industrial and military systems [1], [2].

A sub-family of MCs is called as Indirect MC (IMC) having same advantages together with easier control and safe commutation and it was proposed in [3]. Main objectives in control of IMC is to obtain sinusoidal output currents and sinusoidal input currents with unity power factor. Modulation techniques, control strategies and the digital implementation of the PWM switching patterns were first research attentions on IMC [4]–[6]. Another research perspective on IMC was the reduction of the number of semiconductor switches used, and the family of sparse matrix converter was proposed [7].

In order to avoid relatively complex modulation techniques that can achieve unity power factor at input side and generate sinusoidal waveforms at load side, model predictive control technique can be a good candidate since it has simpler concept, better dynamic response, and easier implementation thanks to recent advancements in digital signal processor technologies [8], [9]. In [10]–[12] a model predictive control approach, whose cost function employs errors in output currents in $\alpha\beta$ frame and input reactive power to meet power-factor performance objective, is proposed. The use of weighting factors in a tailored cost function is unavoidable since it contains different control objectives. In [10], the cost function is updated to include active damping term in order to alleviate the resonance phenomenon of the lightly damped input LC filter. The high-pass filtered active damping current term calculated in the dq synchronous frame from the filter capacitor voltages is somewhat scaled as an output current variable and then subtracted from the output current reference. As a result, the active damping term is indirectly included in the control scheme despite the cost of extra calculation for transformation. The best switching combination among the 24 switching states that gives positive voltage for dc link is selected and applied to converter for next sampling period. A dead time compensation approach is also taken into account to minimize the adverse effect of relatively late updated control signals due to long-lasting calculations [12]. In [13], same model predictive approach as in [10]–[12] is proposed for

four-leg IMC which is used to obtain secure operation for unbalanced load conditions. The predictive approach for four-leg IMC has 48 switching states to be searched.

In [14], in addition to conventional predictive approach for MCs, a cost term is added to cost function to reduce the number of commutations and consequently it decreases the switching losses. In [15], a method to decrease the computational burden of model predictive approach is proposed. This technique eliminates the prediction of three-phase currents at sampling instant $(k+1)$, and reduces the number of reactive power predictions and cost function evaluation from 25 to 10.

In [16]–[18], imposed source current model predictive control is another approach which is used in grid-connected topologies to simultaneously control the input power demands and power factor objective. In this technique, grid side of the power stage is forced to have sinusoidal source currents synchronous or asynchronous according to the power factor demand. The complicated part of this approach is to generate sinusoidal current references which generally require additional PLL environments in hardware or software. These references and source current predictions based on input filter model are then employed in cost function to find best switching combination.

In [2], [19]–[21], predictive control strategies operating at fixed switching frequency of IMC are proposed in order to prevent resonances in the input filter of the matrix converter which is the result of variable switching frequency due to the use of conventional model predictive control. The technique is called as modulated MPC (M2PC) and the cost function is designed to choose optimal adjacent vectors and their corresponding duty cycles. It basically emulates the space vector modulation (SVM) using MPC. In [21], two methods were proposed. First one is the combination of predictive current control with reactive power minimization and active damping technique whereas the second one imposes a sinusoidal source current synchronized with grid voltage to minimize input instantaneous reactive power. The use of fixed switching frequency operation combined with these methods further enhance the performance on the mitigation of input filter resonance. In [20], an indirect approach for M2PC, where two separate cost functions are evaluated for rectifier and inverter sides individually and the use of weighting factors are avoided, is adopted whereas a single cost function to control the IMC is evaluated in [10]–[12].

Predictive approach based on direct power control which only employs active and reactive power errors in cost function can be used in various three-phase power converters without requiring sector information for related control variable in order to synthesize sinusoidal waveform with desired power factor [22]–[24]. In [9], a predictive power control of IMC with active damping is proposed for only resistive load conditions. The errors in load active power, load reactive power, and input reactive power at the instant of $(k+2)$ th are employed in cost function by considering delay compensation. The use of weighting factor is avoided since the cost function employs objectives which have equal importance on the

performance. The active damping technique is introduced to the control scheme by injecting related damping power terms to the load active power reference and input reactive power reference. In [22], Direct Power Control (DPC) with two-step prediction to compensate the control delay is used for active front end rectifier. A PI controller is used to generate input active power reference from the error of DC link voltage. In [23], a similar DPC approach is also applied to current source rectifier used in a battery charge application from three-phase grid. An active damping technique to mitigate input filter resonance is included in the control with a simple novel approach where high-pass filtered damping power term is injected into input active power reference.

In this study, a model predictive control approach with predictive active damping method is proposed to control IMC. No weighting factor is used for the reason that both error terms introduced in the cost function are in the same nature. For the rectifier stage, predictive approach based on imposed sinusoidal source currents is adopted and a novel active damping method, without requiring any pre-calculation of optimal fictitious damping resistor value, is included in the control of rectifier side. For inverter stage, the conventional predictive current control is used. The proposed control scheme is tested by using Matlab/Simulink and collected simulation data show that proposed control method provides reliable operation. Since MPC method has repetitive operations, it is well-suited to implement on a Field-Programmable Gate Arrays (FPGAs) which has pipelining capability. A scheme that shows how the algorithm can be implemented on an FPGA device is also presented.

When this study is compared with the similar counterparts from literature that focus onto any grid connected topology such as IMC and ac-dc current source rectifier with MPC approach, the main contribution of this paper is that, the proposed control method uses a unique direct technique which brings less extra computational burden for reducing THD at the input current. Against to that, [2], [19]–[21] propose modulated MPC emulating the SVM to suppress the input current pollution, and it increases the computational complexity since it requires the selection of optimal adjacent vectors and their corresponding duty cycles. When the computational burden increases, the sampling period increases, and as a direct consequence of that the available bandwidth decreases. In fact, the M2PC used in these studies does not include active damping. It only tries to bypass the problem by keeping the switching frequency away from the resonance frequency of the input filter. However, if any harmonic close to resonance frequency interferes into the system, the M2PC cannot damp it. The use of fixed switching frequency may also increase the switching losses when it is higher than the required, whereas MPC has variable switching frequency according to the operating conditions. While the techniques presented in [10], [23], [25] requires the selection of optimal value for fictitious damping resistor whereas the proposed damping technique in this study does not. In [10], [26], the damping current term is somehow translated into an output variable, and damping effect is obtained via the output

variables in an indirect manner by increasing the computation burden.

The outline of the paper is as follows. The IMC topology and system model are given in section 2. Complete control scheme of MPC is described in section 3. In section 4, reference generation block for supply current peak value is detailed and PI compensator design steps are explained. The details of novel active damping method are presented in section 5. An implementation scheme on how the proposed control is paralleled on an FPGA is provided in section 6. Simulation waveforms and several comparison results in terms of the important evaluation criteria, Total Harmonic Distortion (THD), under different operating conditions of supply and load current are provided in section 7. The study is concluded with section 8.

II. SYSTEM AND PREDICTION MODELS

A. Indirect Matrix Converter Topology

The IMC, shown in Fig. 1, is composed of bidirectional current source rectifier (CSR) and voltage source inverter (VSI) stages. The rectifier side is connected to the grid via a LC filter to prevent the interference of unwanted frequency harmonics to the utility and contains six bidirectional switches. The inverter stage is a full-bridge three-phase circuit whose output is connected to a three-phase RL load.

B. Rectifier Model

The current source rectifier has 9 switching states and three of them gives zero output voltage ($v_{DC}=0$), and another three of them provides positive dc link voltage ($v_{DC}>0$). Since the inverter side requires positive dc link voltage, only three switching combinations among nine switching states can provide positive dc voltage at any sampling instant. Valid switching states, corresponding input and output variables for rectifier side are provided in Table I.

The dc link voltage, v_{DC} , and input current of the rectifier, \mathbf{i}_i , are defined by (1) and (2).

$$v_{DC} = [S_{r1} - S_{r2} \quad S_{r3} - S_{r4} \quad S_{r5} - S_{r6}] \begin{bmatrix} v_{ia} \\ v_{ib} \\ v_{ic} \end{bmatrix} \tag{1}$$

$$= \mathbf{T}_{CSR} \mathbf{v}_i$$

and

$$\begin{bmatrix} i_{ia} \\ i_{ib} \\ i_{ic} \end{bmatrix} = [S_{r1} - S_{r2} \quad S_{r3} - S_{r4} \quad S_{r5} - S_{r6}]^T i_{DC} \tag{2}$$

$$\mathbf{i}_i = \mathbf{T}_{CSR}^T i_{DC}$$

where \mathbf{T}_{CSR} is the transition matrix between input and output of CSR.

C. Inverter Model

The voltage source inverter has 8 valid switching vectors, and two of them are zero vectors. Valid switching states and corresponding input and output variables for inverter side are provided in Table II.

TABLE I
RECTIFIER SIDE'S VALID SWITCHING STATES

#	S_{r1}	S_{r2}	S_{r3}	S_{r4}	S_{r5}	S_{r6}	i_{sa}	i_{sb}	i_{sc}	VDC
1	1	0	0	0	0	1	i_{DC}	0	$-i_{DC}$	v_{ac}
2	0	0	1	0	0	1	0	i_{DC}	$-i_{DC}$	v_{bc}
3	0	1	1	0	0	0	$-i_{DC}$	i_{DC}	0	$-v_{ab}$
4	0	1	0	0	1	0	$-i_{DC}$	0	i_{DC}	$-v_{ac}$
5	0	0	0	1	1	0	0	$-i_{DC}$	i_{DC}	$-v_{bc}$
6	1	0	0	1	0	0	i_{DC}	$-i_{DC}$	0	v_{ab}

TABLE II
INVERTER SIDE'S VALID SWITCHING STATES

#	S_{i1}	S_{i2}	S_{i3}	S_{i4}	S_{i5}	S_{i6}	v_{uv}	v_{vw}	v_{wu}	i_{DC}
1	1	0	0	1	0	1	v_{DC}	0	$-v_{DC}$	i_{ou}
2	1	0	1	0	0	1	0	v_{DC}	$-v_{DC}$	$i_{ou} + i_{ov}$
3	0	1	1	0	0	1	$-v_{DC}$	v_{DC}	0	i_{ov}
4	0	1	1	0	1	0	$-v_{DC}$	0	v_{DC}	$i_{ov} + i_{ow}$
5	0	1	0	1	1	0	0	$-v_{DC}$	v_{DC}	i_{ow}
6	1	0	0	1	1	0	v_{DC}	$-v_{DC}$	0	$i_{ou} + i_{ow}$
7	1	0	1	0	1	0	0	0	0	0
8	0	1	0	1	0	1	0	0	0	0

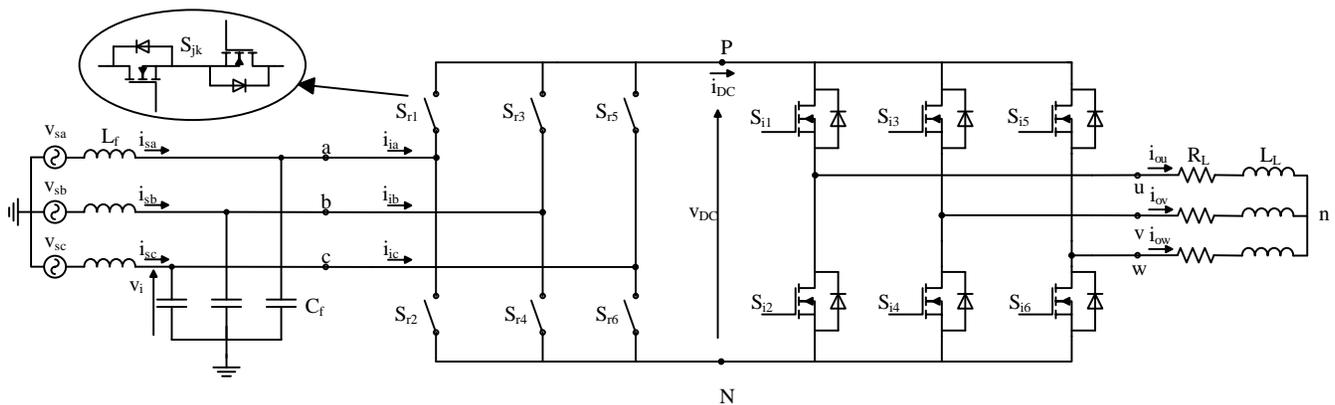


Fig.1. Indirect Matrix Converter

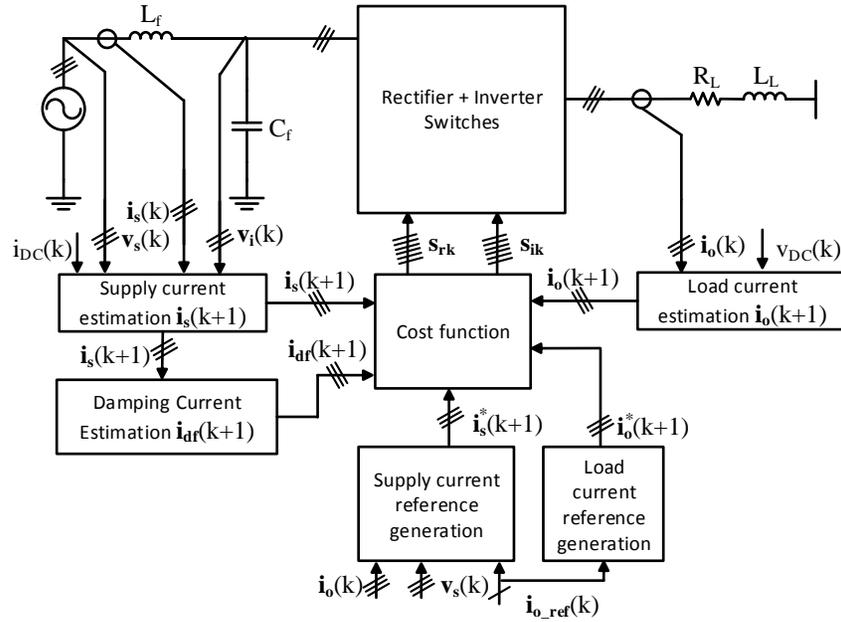


Fig.2. Model Predictive Control Scheme for IMC

The relationship between input and output current and voltage variables of VSI is defined in (3) and (4).

$$\begin{aligned} \mathbf{i}_{DC} &= \begin{bmatrix} S_{i1} & S_{i3} & S_{i5} \end{bmatrix} \begin{bmatrix} i_{ou} \\ i_{ov} \\ i_{ow} \end{bmatrix} \\ &= \mathbf{T}_{VSI} \mathbf{i}_o \end{aligned} \quad (3)$$

$$\begin{aligned} \begin{bmatrix} v_{oun} \\ v_{ovn} \\ v_{own} \end{bmatrix} &= \frac{v_{DC}}{3} \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} \begin{bmatrix} S_{i1} & S_{i3} & S_{i5} \end{bmatrix}^T \\ \mathbf{v}_o &= \frac{v_{DC}}{3} \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} \mathbf{T}_{VSI}^T \end{aligned} \quad (4)$$

D. Prediction Model for Rectifier Side

In order to compensate the supply current error, the supply current must be predicted. In this regard, unity power factor is achieved. Input filter model of CSR is used to predict the supply current value at next sampling interval and the discrete model is defined in (5).

$$\begin{bmatrix} \mathbf{v}_i(k+1) \\ \mathbf{i}_s(k+1) \end{bmatrix} = \Phi \begin{bmatrix} \mathbf{v}_i(k) \\ \mathbf{i}_s(k) \end{bmatrix} + \Gamma \begin{bmatrix} \mathbf{v}_s(k) \\ \mathbf{i}_i(k) \end{bmatrix} \quad (5)$$

Supply current prediction, which will be used in the cost function, is based on (6).

$$\begin{aligned} \mathbf{i}_s(k+1) &= \Phi(2,1)\mathbf{v}_i(k) + \Phi(2,2)\mathbf{i}_s(k) \\ &+ \Gamma(2,1)\mathbf{v}_s(k) + \Gamma(2,2)\mathbf{i}_i(k) \end{aligned} \quad (6)$$

E. Prediction Model for Inverter Side

In order to control the output current, the discrete time model of load is required, and it is given by (7).

$$\mathbf{i}_o(k+1) = \mathbf{i}_o(k) \left[1 - \frac{R_L T_s}{L_L} \right] + \frac{T_s}{L_L} \mathbf{v}_o(k) \quad (7)$$

III. MODEL PREDICTIVE CONTROL

The schematic illustration of the proposed MPC technique is depicted in Fig. 2. For the k^{th} sampling instant supply voltages \mathbf{v}_s , supply currents \mathbf{i}_s , input filter capacitor voltages \mathbf{v}_i and output load currents \mathbf{i}_o are measured. The dc link voltage is predicted based on (1), and it is calculated for each switching vector of CSR side using input filter capacitor voltage measurements at k^{th} instant. If this predicted dc link voltage is positive, then related rectifier switching vector is evaluated for each switching combination of inverter side. Firstly, load currents are predicted based on (7) for each inverter switching vector by substituting the predicted dc link voltage. These load current estimations are used for cost function evaluation in (8). Secondly, the dc link current is predicted based on (3) for each inverter switching vector and then this dc link current prediction for related inverter switching vector is used to predict supply currents based on (6) for related rectifier switching combination. Then these supply current predictions evaluated in the cost function defined by (9). These two cost functions are combined into a single cost function as shown in (10).

$$g_1 = \sum_{j=u,v,w} |i_{oj}^*(k+1) - i_{oj}(k+1)|^2 \quad (8)$$

$$g_2 = \sum_{j=a,b,c} |i_{sj}^*(k+1) - i_{sj}(k+1)|^2 \quad (9)$$

$$g_{tot} = g_1 + g_2 \quad (10)$$

Once the cost function evaluation process is over for the inverter stage, whole operation is repeated for the next candidate dc-link voltage generated from rectifier stage. Basically, cascaded structure is considered for proper operation. The rectifier stage has 3 valid switching states that give positive dc link voltage and the inverter stage has 8 switching states and there are totally 24 switching combinations in the single step prediction horizon. The switching combination that provides minimum cost function value is picked and applied to IMC for the next time interval. The cost function includes current terms for supply and load sides and as a result no weighting factor is required to distinguish the control importance one over another.

IV. REFERENCE GENERATION AND COMPENSATOR DESIGN

The approach to generate supply currents synchronized with grid voltages is illustrated in Fig. 3 where $i_{o_ref}(k)$ is the magnitude of load current reference. Measured three phase load currents are firstly transformed to dq frame and then the magnitude of rotating space vector is calculated. The error signal, which is difference between reference load current and measured load current, is fed into PI compensator and a constant multiplication factor, designated with m , is generated. This constant is then used to generate three phase supply current references by multiplying it with instantaneous grid voltages.

In order to design the controller, dynamic model of the plant is required, and it is derived by equating input power to output power with the assumption of a lossless converter. Another assumption in this derivation is that input filter has much faster dynamics than the load dynamics and hence its effect is neglected in plant modeling. Accordingly, (11) is obtained.

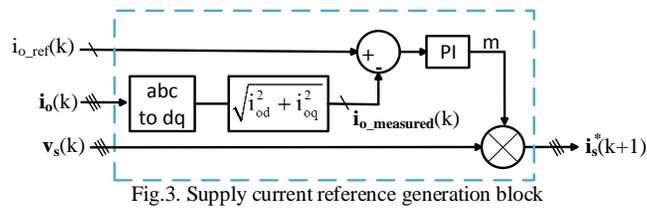


Fig.3. Supply current reference generation block

$$1.5V_s I_s = I_o^2 (R_L + sL_L) \quad (11)$$

After usual perturbation and linearization steps on (11), dynamic model can be expressed by (12). The open loop bode plot of dynamic model with the parameters; $I_o=5A$, $V_s=311V$, $R_L=10\Omega$ and $L_L=10mH$ is given in Fig. 4.

$$\frac{\dot{i}_o}{\dot{V}_s} = \frac{1.5V_s}{2I_o (R_L + sL_L)} \quad (12)$$

A PI controller is designed by setting the cross-over frequency at 300 Hz and the phase margin to 67°. Bode plot of compensated loop gain is shown in Fig. 5. Frequency domain equation of PI compensator is given in (13) and discrete time

domain equivalent, which is obtained by applying Forward Euler discretization, is given in (14).

$$C(s) = k_p + \frac{k_i}{s} \quad (13)$$

$$u(k) = u(k-1) + k_p e(k) + (k_i T_s - k_p) e(k-1) \quad (14)$$

where u is the output of the PI compensator and e is the input of PI compensator and the error in the control variable.

V. ACTIVE DAMPING METHOD

Lightly damped input LC filter of CSR causes unwanted harmonics generation. In particular, this issue appears when system harmonics around the resonant frequency from utility or rectifier itself is interfered to the filter. These harmonics pollute the grid currents and the harmonic level in the grid current can become unacceptable. A usual way to damp the resonance of LC filter is to add a resistor in parallel to filter capacitor at the expense of a drastic reduction in the efficiency. The function of this parallel resistor can be emulated by means of a technique called as active damping employed in the control scheme, and a damping effect can be obtained without disturbing the efficiency.

In this study a novel active damping method is proposed, and its conceptual scheme is illustrated in Fig. 6. The aim in active damping is to decrease the supply current harmonic contents. Thus, three phase supply currents can be directly

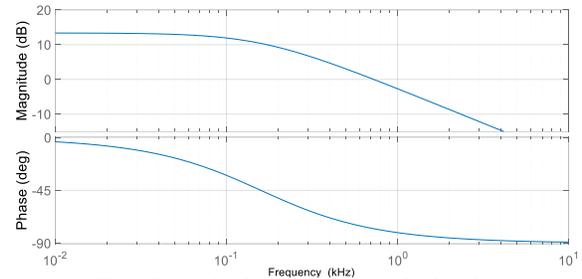


Fig.4. Bode plot for open loop transfer function

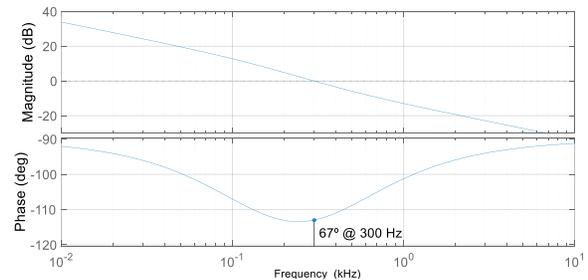


Fig.5. Bode plot for compensated loop function



Fig.6. Damping current estimation scheme

used in damping scheme instead of using input filter capacitor voltage with a suitable damping resistance value as have been widely adopted in literature [10], [23], [25]. Since the MPC technique has a predictive nature, predicted value of supply current $i_s(k+1)$ for next sampling interval is used in active

damping. These three phase currents include components located at fundamental grid frequency and higher frequencies. A High Pass Filter (HPF), whose transfer functions in continuous and discrete time domain are defined in (15) and (16) respectively, can be used to remove fundamental component and to obtain high frequency components which are responsible of supply current pollution. If these three phase currents $i_{df}(k+1)$ composed of high frequency components are subtracted from the supply current references, then this effectively reduces the magnitudes of high frequency harmonics in supply current and filter capacitor voltage. Accordingly, the cost function g_2 must be updated as shown in (17).

The design of HPF is essential in terms of filtering capability of high-frequency components to obtain a good damping effect. In order to alleviate fundamental component and allow to pass higher order harmonics, a first order HPF with a bandwidth of 500 Hz is used to generate filtered version of supply current predictions, $i_{df}(k+1)$.

$$HPF(s) = \frac{3.1831e-04s}{1 + 3.1831e-04s} \quad (15)$$

$$i_{df}(k+1) = 0.9372 i_{df}(k) + i_d(k+1) - i_d(k) \quad (16)$$

$$g_2 = \sum_{j=a,b,c} |i_{sj}^*(k+1) - i_{dfj}(k+1) - i_{sj}(k+1)|^2 \quad (17)$$

VI. FPGA IMPLEMENTATION SCHEME

A big disadvantage of model predictive control method is high sampling rate required to achieve robust operation. MPC contains large amount of control calculations and this causes

an issue in real-time implementation. To handle with high computational burden imposed by MPC algorithm, FPGA devices can be used to perform all control calculations in parallel. Since MPC contains independent math calculations for different switching combinations, these calculations can be implemented simultaneously. The main principle of FPGA-based MPC is that independent calculation blocks are concurrently performed in different FPGA areas by dividing the instructions processing cycle into distinct stages of processing. This significantly increases total bandwidth of the controller and whole control algorithm can be performed in a shorter time. The main idea of how paralleling the control calculation of the proposed method is shown in Fig. 7. The first step is to extract the required measurement from sensing board and these measurements are used to identify which rectifier switching combination provide dc-link voltage to the inverter stage. Second step is to move proper rectifier switching combination information to the buffer. After that, output load current is predicted for different inverter switching states. As can be seen from Fig. 7, instead of doing serial computation, independent prediction blocks are distributed to the different areas of the FPGA. This reduces the total execution time and provides higher control bandwidth. The next step is to predict the supply current for different selected rectifier state. The last step is to perform exhaustive search so as to obtain optimum control action for the next time interval.

This pipelining procedure can be implemented in cost-effective FPGA solutions, such as DEO-Nano Board from Terasic. The aforementioned board has Cyclone IV and 50 MHz clock speed. Since FPGA does not have fixed core

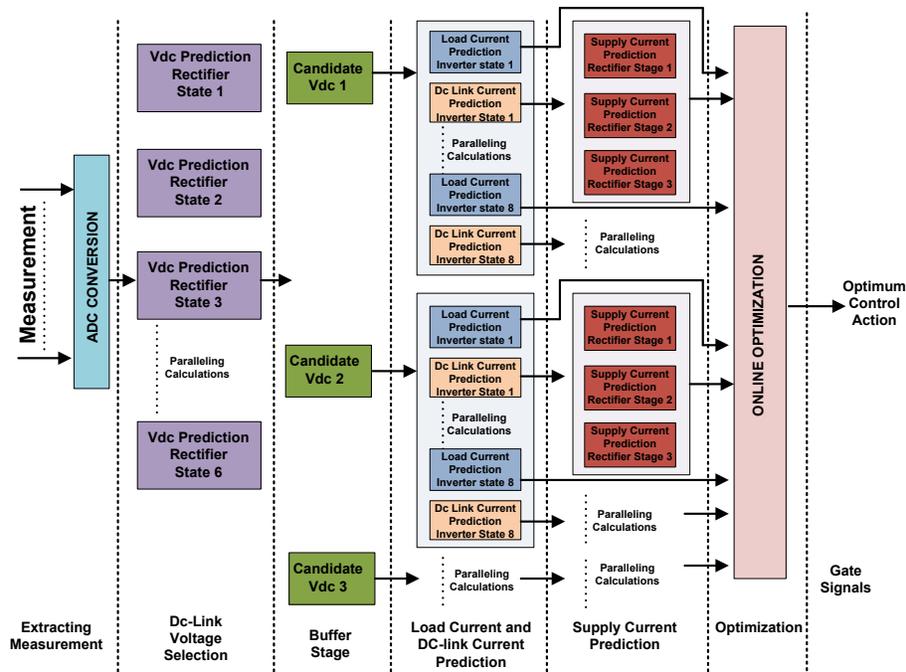


Fig.7. Parallel implementation scheme of the proposed control on a FPGA device

TABLE III
APPROXIMATE TOTAL EXECUTION TIME FOR EACH STEP

Tasks	Clock Cycle	Execution Time (μ s)
Extracting Measurement	4	0.08
DC-Link Voltage Selection	2	0.04
Buffer Stage	1	0.02
Load Current Prediction and Dc-Link Current Prediction	15	0.30
Supply Current Prediction	11	0.22
Cost Calculation	28	0.56
Optimization	37	0.74
Total	98	1.96

architecture, total performance can be calculated Clock Speed \times Core/(1 Clock Cycle). The effective core number depends on the pipelining stage that designed by the user. Approximate total execution time for each task is tabulated in Table III in case of fully pipelining architecture. Since data dependency occurs due to cascaded MPC block, there is a limit in decreasing the running time of proposed MPC algorithm. The values reported in Table III are extracted by performing VHDL design simulation using ModelSim ALTERA. The approximated values are measured from ModelSim ALTERA scope and state-machine cycles are observed by using the simulation tools.

VII. SIMULATION RESULTS

The proposed control scheme is validated by simulation work performed in Matlab/Simulink with parameters tabulated in Table IV.

Waveforms presented in Fig. 8 are obtained when a load current reference with a peak value of 10 A at 100 Hz is applied to the converter. Fig. 8a shows the supply voltage and current (x50) for phase-a and as it is observed from the figure that almost unity power factor for the grid side is achieved with the proposed controller. Three-phase supply current references generated by the method depicted in Fig. 3 are shown in Fig. 8b. They are in phase with grid voltages as it is observed in Fig. 8a and supply currents track sinusoidal references as shown in Fig. 8c. Load current references and measurements are shown in Fig. 8d and load currents show a good performance in terms of reference tracking. Power components of supply side are also shown in Fig. 8e and reactive power component remains at zero as it is expected in unity power factor control. From Fig. 8 it can be concluded

TABLE IV
SIMULATION PARAMETERS

Parameter	Description	Values
T_s	Sampling Period	20 μ s
R_L	Load Resistance	10 Ω
L_L	Load Inductance	10 mH
L_f	Input Inductance	400 μ H
C_f	Input Capacitance	21 μ F
R_f	Inp. Filt. Damp Resistance	0.5 Ω
V_s	Supply Peak Voltage	311 V
f_g	Supply frequency	50 Hz

that the proposed control gives enough flexibility to control the relatively complex IMC topology in terms of input power factor correction an output load current control.

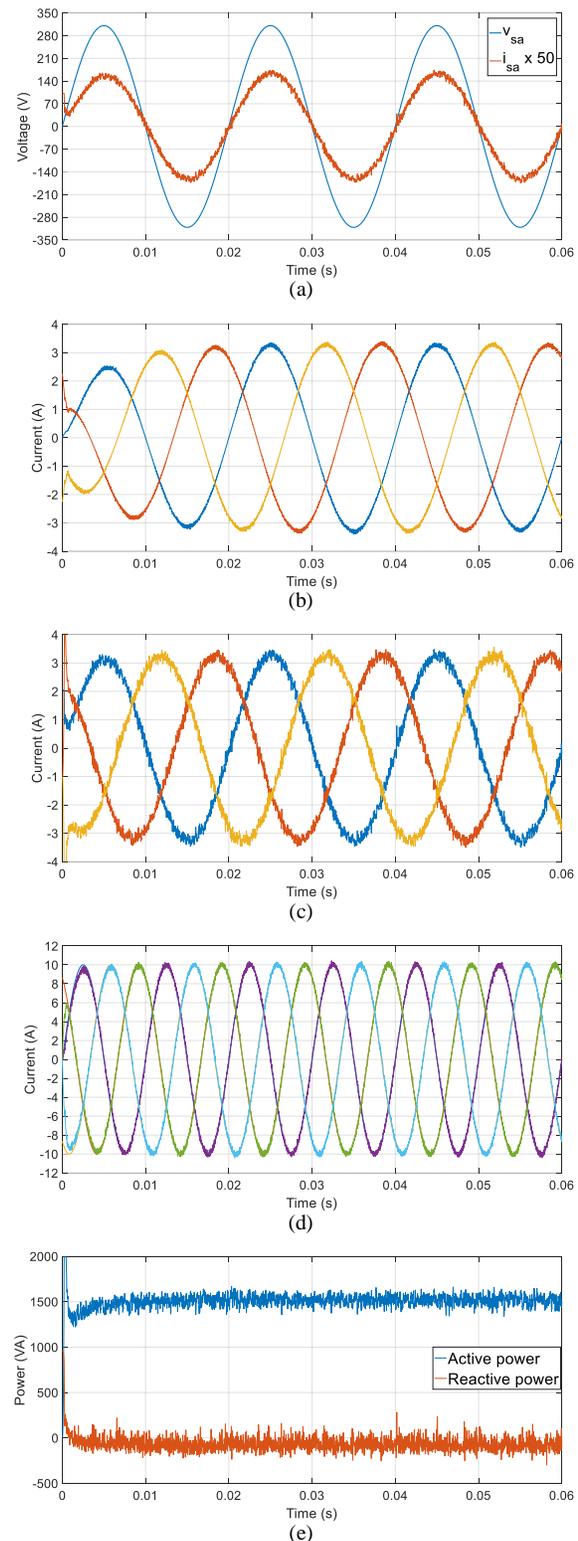


Fig.8. Waveforms for $I_{o_REF}=10A @ 100 Hz$; (a) Supply voltage and current (x50) for phase-a, (b) Three phase supply current references, (c) Three phase supply currents, (d) Three phase load current references and load currents, (e) Supply power components

TABLE V
THD RESULTS OBTAINED UNDER DIFFERENT CONDITIONS
APPROXIMATE

Sampling and load reference ↓	$I_{O_REF} \rightarrow$	5A		10A	
	THD \rightarrow	load	supply	load	supply
	Control ↓	5A	0.84A	10A	3.26A
$T_s=20\mu s$ Load ref. at 50Hz	without damping	3.03%	30.02%	1.59%	7.58%
	with damping	3.32%	16.21%	1.97%	5.46%
$T_s=20\mu s$ Load ref. at 100Hz	without damping	2.90%	33.36%	1.63%	7.76%
	with damping	3.28%	15.23%	2.01%	5.58%
$T_s=50\mu s$ Load ref. at 50Hz	without damping	8.28%	62.62%	5.22%	28.24%
	with damping	9.94%	42.31%	6.76%	23.88%

$K_p=0.288, K_i=669.56$

Table V compares the THD levels in supply current for different load current reference demands. The proposed controller achieves a good performance for THD levels in load currents with a sampling period of 20 μs . The THD in supply currents is generally problematic in the control of IMC especially at low current demands. For supply current peak value of 0.84 A, the THD is 30.02% and it reduces to 16.21% with the inclusion of proposed damping technique. For higher current level of supply, the THD value inherently decreases and active damping also reduces these THD values by around 2%. Activating the damping term in the proposed control causes a slight increment in THD of load currents, but it is in the range of acceptable levels.

Fig. 9 compares the supply current FFT analysis results with and without active damping for the case where reference load current of 5A at 100 Hz is applied to converter, and as a result

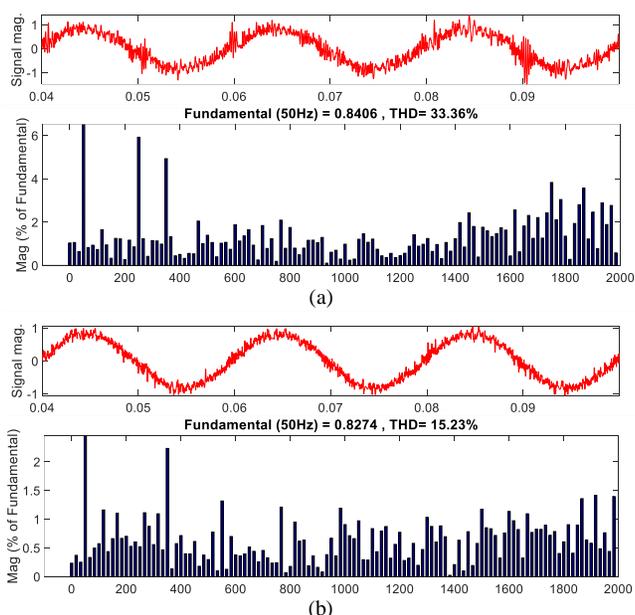


Fig.9. FFT analysis for supply current for $I_{o_REF}=5A$ @ 100 Hz; (a) without active damping, THD=33.36% (b) with active damping, THD=15.23%.

there is a supply current having a 0.84A peak value at 50 Hz. Inclusion of the active damping shows a considerable reduction in THD values especially for low current levels.

VIII. CONCLUSION

This paper proposes a model predictive approach to control IMC topology. Following two goals are simultaneously achieved; unity power factor with relatively low THD for supply current and sinusoidal current with desired magnitude and frequency for three phase RL load. Magnitude of supply current reference is generated from the error in the magnitude of load current space vector by a PI compensator. Input filter model is used to predict future behavior of the supply current so as to control the rectifier stage. These current predictions in abc frame are then employed in the cost function. Active damping current term, which is obtained by high-pass filtering the supply current prediction, is also included into cost function of rectifier side to mitigate resonance phenomenon of input LC filter. Adding this damping term in the cost function helps to reduce harmonics in supply current. Conventional MPC approach, whose cost function employs load currents in abc frame, is adopted for the inverter side to synthesize sinusoidal current with desired peak value and frequency. In order to find best switching combination, a typical exhaustive search algorithm is performed for 24 different switching combinations. The one among these combinations minimizes the cost function is selected.

REFERENCES

- [1] J. W. Kolar, M. Baumann, F. Schafmeister, and H. Ertl, "Novel three-phase AC-DC-AC sparse matrix converter," pp. 777–791, 2003.
- [2] M. Rivera, L. Tarisciotti, P. Wheeler, and P. Zanchetta, "Predictive control of an indirect matrix converter operating at fixed switching frequency and without weighting factors," *IEEE Int. Symp. Ind. Electron.*, vol. 2015-Sept, pp. 1027–1033, 2015.
- [3] K. Iimori, K. Shinohara, O. Tarumi, Zixun Fu, and M. Muroya, "New current-controlled PWM rectifier-voltage source inverter without DC link components," pp. 783–786, 2002.
- [4] J. Rodriguez, M. Rivera, J. W. Kolar, and P. W. Wheeler, "A review of control and modulation methods for matrix converters," *IEEE Trans. Ind. Electron.*, vol. 59, no. 1, pp. 58–70, 2012.
- [5] C. Klumpner, F. Blaabjerg, I. Boldea, and P. Nielsen, "New modulation method for matrix converters," *IEEE Trans. Ind. Appl.*, vol. 42, no. 3, pp. 797–806, 2006.
- [6] L. Helle, K. B. Larsen, A. H. Jorgensen, S. Munk-Nielsen, and F. Blaabjerg, "Evaluation of Modulation Schemes for Three-Phase to Three-Phase Matrix Converters," *IEEE Trans. Ind. Electron.*, vol. 51, no. 1, pp. 158–171, 2004.
- [7] J. W. Kolar, F. Schafmeister, S. D. Round, and H. Ertl, "Novel three-Phase AC-AC sparse matrix converters," *IEEE Trans. Power Electron.*, vol. 22, no. 5, pp. 1649–1661, 2007.
- [8] S. Vazquez, J. Rodriguez, M. Rivera, L. G. Franquelo, and M. Norambuena, "Model Predictive Control for Power Converters and Drives: Advances and Trends," *IEEE Trans. Ind. Electron.*, vol. 64, no. 2, pp. 935–947, 2017.
- [9] J. Lei *et al.*, "Predictive Power Control of Matrix Converter with Active Damping Function," *IEEE Trans. Ind. Electron.*, vol. 63, no. 7, pp. 4550–4559, 2016.
- [10] M. Rivera, P. Correa, J. Rodriguez, I. Lizama, and J. Espinoza, "Predictive control of the indirect matrix converter with active damping," *2009 IEEE 6th Int. Power Electron. Motion Control Conf. IPEMC '09*, pp. 1738–1744, 2009.
- [11] J. Rodriguez, J. Kolar, J. Espinoza, M. Rivera, and C. Rojas, "Predictive current control with reactive power minimization in an indirect matrix converter," *Proc. IEEE Int. Conf. Ind. Technol.*, pp. 1839–1844, 2010.

- [12] P. Correa, J. Rodríguez, M. Rivera, J. R. Espinoza, and J. W. Kolar, "Predictive control of an indirect matrix converter," *IEEE Trans. Ind. Electron.*, vol. 56, no. 6, pp. 1847–1853, 2009.
- [13] C. F. García, M. E. Rivera, J. R. Rodríguez, P. W. Wheeler, and R. S. Pena, "Predictive Current Control with Instantaneous Reactive Power Minimization for a Four-Leg Indirect Matrix Converter," *IEEE Trans. Ind. Electron.*, vol. 64, no. 2, pp. 922–929, 2017.
- [14] R. Vargas, U. Ammann, and J. Rodríguez, "Predictive approach to increase efficiency and reduce switching losses on matrix converters," *IEEE Trans. Power Electron.*, vol. 24, no. 4, pp. 894–902, 2009.
- [15] M. Siami, D. Arab Khaburi, and J. Rodriguez, "Simplified Finite Control Set-Model Predictive Control for Matrix Converter-Fed PMSM Drives," *IEEE Trans. Power Electron.*, vol. 33, no. 3, pp. 2438–2446, Mar. 2018.
- [16] P. Zavala *et al.*, "Predictive control of a current source rectifier with imposed sinusoidal input currents," *IECON Proc. (Industrial Electron. Conf.)*, pp. 5842–5847, 2013.
- [17] M. Rivera, L. Tarisciotti, and P. Wheeler, "Indirect model predictive control with imposed sinusoidal source currents for a Direct Matrix Converter Working at fixed switching frequency," *Proc. - 2017 IEEE South. Power Electron. Conf. SPEC 2017*, vol. 2018-Janua, pp. 1–6, 2018.
- [18] M. Gokdag and O. Gulbudak, "Model predictive control of AC-DC matrix converter with unity input power factor," *Proc. - 2018 IEEE 12th Int. Conf. Compat. Power Electron. Power Eng. CPE-POWERENG 2018*, pp. 1–5, 2018.
- [19] M. Rivera, C. Uribe, L. Tarisciotti, P. Wheeler, and P. Zanchetta, "Predictive control of an indirect matrix converter operating at fixed switching frequency and unbalanced AC-supply," *Proc. - 2015 IEEE Int. Symp. Predict. Control Electr. Drives Power Electron. Preced. 2015*, pp. 38–43, 2016.
- [20] M. Rivera, L. Tarisciotti, P. Wheeler, and P. Zanchetta, "Predictive control of an indirect matrix converter operating at fixed switching frequency," *Int. Conf. Power Eng. Energy Electr. Drives*, vol. 2015-Sept, pp. 635–640, 2015.
- [21] M. Rivera, M. Amirbande, A. Vahedi, L. Tarisciotti, and P. Wheeler, "Predictive control strategies operating at fixed switching frequency for input filter resonance mitigation in an indirect matrix converter," *Proc. - 2017 IEEE South. Power Electron. Conf. SPEC 2017*, vol. 2018-Janua, pp. 1–6, 2018.
- [22] P. Cortés, J. Rodríguez, P. Antoniewicz, and M. Kazmierkowski, "Direct power control of an AFE using predictive control," *IEEE Trans. Power Electron.*, vol. 23, no. 5, pp. 2516–2523, 2008.
- [23] M. Gokdag and O. Gulbudak, "Model Predictive Control for Battery Charger Applications with Active Damping," in *2019 1st Global Power, Energy and Communication Conference (GPECOM)*, 2019, pp. 140–145.
- [24] S. Vazquez, A. Marquez, R. Aguilera, D. Quevedo, J. I. Leon, and L. G. Franquelo, "Predictive Optimal Switching Sequence Direct Power Control for Grid-Connected Power Converters," *IEEE Trans. Ind. Electron.*, vol. 62, no. 4, pp. 2010–2020, 2015.
- [25] M. Gökdağ and O. Gülbudak, "Imposed Source Current Predictive Control for Battery Charger Applications with Active Damping," *Sak. Univ. J. Sci.*, pp. 964–971, Oct. 2019.
- [26] P. Correa, J. Rodriguez, I. Lizama, and D. Andler, "A Predictive Control Scheme for Current-Source Rectifiers," *IEEE Trans. Ind. Electron.*, vol. 56, no. 5, pp. 1813–1815, May 2009.

BIOGRAPHIES



MUSTAFA GOKDAG received the B.Sc. degree with Honor in electrical and electronics engineering from Firat University, Turkey, in 2009, and the M.Sc. and Ph.D. degrees in electrical and electronics engineering from Karabuk University, Turkey, in 2011 and 2016 respectively. From 2009 to 2016, he was a Research Assistant with

the department of electrical and electronics engineering in Karabuk University. Since 2016, he has been an Assistant Professor in same department. His research interests include modeling and control of dc-dc power converters and model predictive control of ac-dc, dc-ac, and ac-ac power converters for renewable and electrical drives.



OZAN GULBUDAK received the B.Sc. and the M.Sc. degree in electrical engineering from Mersin University, Turkey in 2008 and 2010. He received Ph.D. degree from the University of South Carolina, Columbia, USA. Since 2017, he has been with Karabuk University, where is currently Assistant Professor. His

research interests include model predictive control, development of control platforms based on FPGA, direct matrix converters, inverter topologies and motor drives.

The Impact of the Government's Incentives on Increasing Investment in Turkey's Solar Photovoltaic Power Plants

F.ERKEN

Abstract— Exhibiting a high economic growth rate in recent years, energy demand has increased rapidly in Turkey. Despite having a great potential in terms of renewable energy sources, Turkey has not utilized this potential sufficiently to meet this demand. However, among various renewable energy sources, the solar energy stands out in terms of investments and incentives in Turkey. In this study, the progress and the outlook of the renewable energy, especially photovoltaic systems in Turkey are reviewed considering the potential, the incentives and the investments. In this context, ratio of the Photovoltaic Solar Power Plants within both overall installed power and the total generated energy over the years is investigated geopolitically. Ongoing and planned investments for Photovoltaic Solar Power Plants and the incentives provided by the government are reviewed by considering the geographical regions of the country. It is observed that the installed power capacity of active solar photovoltaic power plants increases by more than four times compared to the previous year, as of the end of 2017, owing to incentives and investments. It is expected that the solar energy power capacity among the renewable energy sources contributes significantly to the total installed power capacity over the next few years in Turkey.

Index Terms— Renewable energy sources, solar energy, solar photovoltaic power plants, incentives and investments, energy projection.

I. INTRODUCTION

THE GLOBAL energy demand has increased substantially in the last few decades owing to the many technological developments [1]. Today, fossil fuels are the primary energy source. However, their limited availability and negative environmental impact has prompted countries to pursue alternative energy sources [2, 3]. Environmental concerns and the urgent need for environmental protection are increasing rapidly as a result of supplying the global energy demand produced from the fossil-based oil, coal and gas etc. The utilization of renewable energy for environmentally friendly energy systems has alternatively increased interests [4, 5]

FARUK ERKEN, is with Department of Electrical and Electronics Engineering of Kastamonu University, Kastamonu, Turkey, (e-mail: ferken@kastamonu.edu.tr).

 <https://orcid.org/0000-0003-2048-1203>

Manuscript received September 27, 2019; accepted December 16, 2019.
DOI: [10.17694/bajece.626145](https://doi.org/10.17694/bajece.626145)

Besides, the quantity of renewable energy studies has increased, and the studies themselves have intensified targeting cheap and clean energy [2]. Particularly in recent years, eco-friendly Renewable Energy Sources (RESs) (also known as alternative energy sources) can be used to regenerate energy. RESs are considered as significant energy resources in many countries all around the world. RESs researches ensure that the world's growing energy demand can be met using clean and economical ways. Finally, safety concerns have increased the importance of RESs such as solar, wind, biomass and geothermal energy, as well as technologies including fuel cells [6-12].

According to the World Bank's World Development Indicators [13], about 67.03% of the total global energy is produced from the burning of fossil-based oil, coal and gas, which produces carbon dioxide (CO₂) emissions, intensifying global warming (through the greenhouse effect), as shown in Fig. 1 [14]. In the near future, billions of people, especially in developing countries, are expected to face many problems due to climate change [15], as global warming is deemed one of the most serious threats in human history [16, 17].

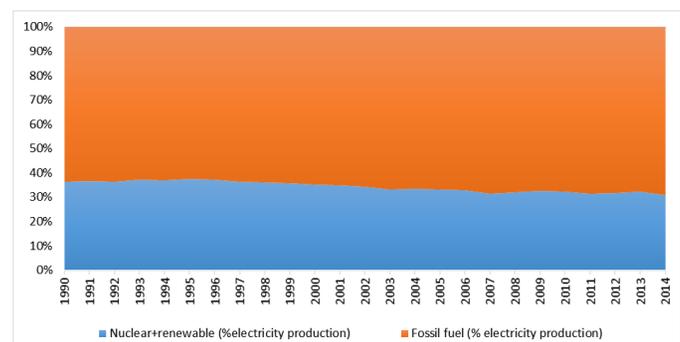


Fig.1. Global electricity production rates measured as the percentage contribution from fossil fuels (coal, oil and gas) and low-carbon sources (nuclear, hydropower, biomass, wind, solar, geothermal and marine power) [11].

Global CO₂ emissions to the atmosphere have increased from 22,112 Mt in 1997, before the Kyoto protocol in December 1997, to 32,294 Mt in 2015, at a rate of 46% [18]. In fact, this rate is expected to increase even further, by almost 40%, according to the Energy Information Administration's International Energy Outlook 2009, during the period 2006–2030 [19]. CO₂ emissions in Turkey have also increased from 176.6 Mt in 1997 to 317.2 Mt in 2015, at a rate of 79.6% [18]. In Fig. 2, the CO₂ emissions trend in Turkey, between 1971

and 2015, are presented as percentage contributions from fossil fuels (coal, oil and gas). The greenhouse gas emissions continue to rise in Turkey. According to Turkish Statistical Institute, this trend has not deteriorated in 2016 and the total emission amount was 496.1 Mt CO₂ equivalent with an increase of 5.6% compared to the previous year [20]. The CO₂ emissions increased to 526.3 Mt in 2017. The results clearly indicate that the CO₂ emissions in Turkey are approximately two times greater than those in the world [18].

A major benefit of using renewable energy is the reduction of pollutant emission in the air, in the form of heating gases [21]. Thus, currently, mankind's aim is to find alternative sources of energy to enable life on the ground, without relying entirely on fossil fuels [22].

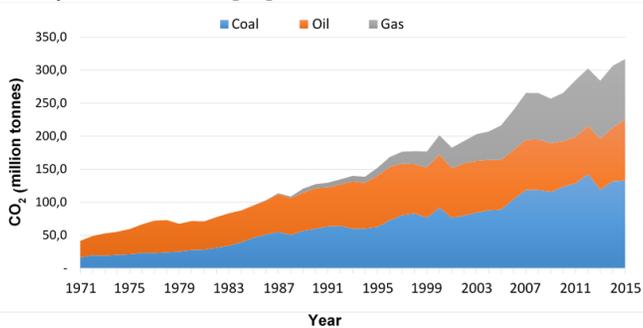


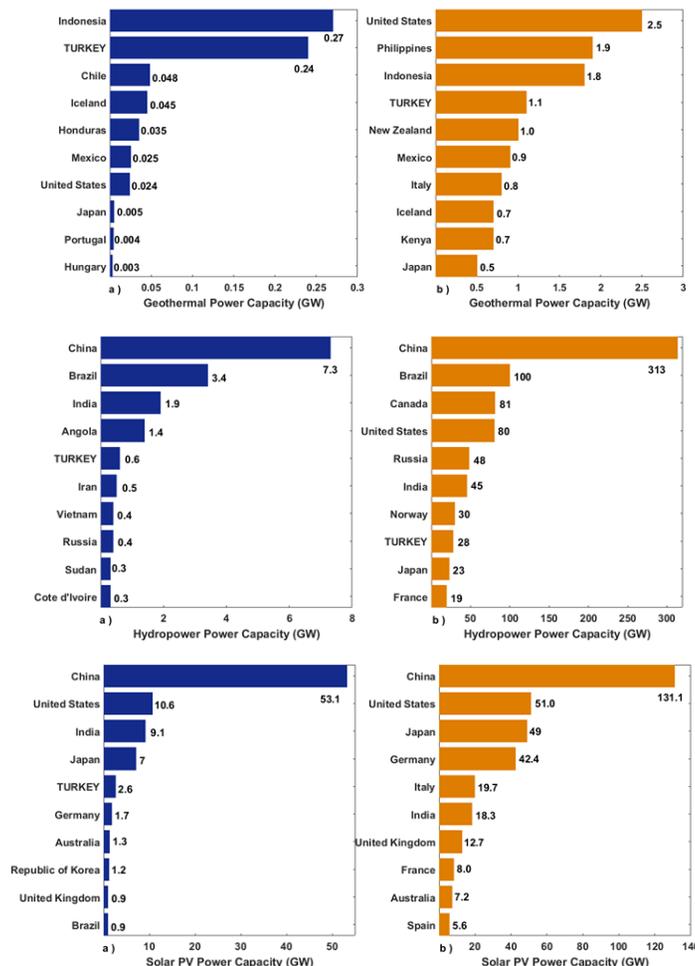
Fig.2. Trend of CO₂ emission of Turkey from fuel combustion between 1971 and 2015 [16].

In today's world, awareness regarding such sources is a necessity to expand the use of all renewable energy sources because environmental problems have increased tremendously, and consequently the adverse effects of climate change have become crucial [16]. However, 26.5% of the global electricity is produced from renewable energy sources with hydropower providing about 16.4%, as 73.5% of electricity production is provided from non-renewable energy sources. Most of countries in the world target the using of the renewable energy sources and regulate their policies and several jurisdictions to increase their existing capacity [23]. As a result of these regulations, global renewable power capacity in 2017 increased to 2,195 GW while it was around 1,000 GW in 2007. Overall, renewable power generating capacity in 2017 reached its greatest annual rate with an increasing 178 GW raising total capacity by almost 9% over 2017. The global hydropower capacity reached to 1,114 GW with an addition of 19 GW despite it has the smallest annual enhancement over the last five years. The leader country in hydropower installed power capacity is China, accounting for nearly 40% of new installations in 2017. Other following countries that added significant capacity are Brazil, India, Angola and Turkey. Wind power is in demand renewable energy sources with more than 52 GW added globally in 2017 compared to 2016. The world's total capacity increased nearly 11%, to 539 GW in 2017. China is the leader country both in total installed and added capacities of wind energy. Turkey is in the top 10 for the new installations in the end of 2017. Geothermal energy is cost-effective, sustainable, and environmentally friendly energy resources. However, drilling and exploration for deep resources is very expensive. Besides, geothermal energy has been limited to areas near tectonic plate boundaries. United

States, The Philippines, Indonesia, Turkey, New Zealand, Mexico, Italy, Iceland, Kenya and Japan are the countries that have the largest amounts of geothermal power generating capacity at the end of 2017. For the new installations, Indonesia and Turkey lead with the rate of 39% and 34% during the year, respectively. Turkey's net additions were at least 243 MW, for a total of 1.1 GW at the end of 2017 [23].

Solar energy, as one of the most important types of RESs, can be used directly in residential and commercial buildings, as well as in the lighting and heating industry. In addition, electricity can be generated through solar energy, using sun-tracking concentrating collectors. Alternatively, electricity can be directly obtained from solar energy using photovoltaic (PV) solar cells that convert photon energy (hv) into electrical energy [2]. Although solar PV power plants do not require mechanical maintenance since they do not have moving parts, their efficiency decreases over the years compared to other renewable power plants.

Turkey has a big potential for solar energy exploitation in the European countries because of its geographic location [27]. The rate of the solar PV in global electricity production is 1.9% in the end of 2017. Turkey provides its electricity generation from solar energy with the rate of 6.82%, which is above the world average (see Table 1) [27 28]. The Fig. 3 shows the new installations and total power capacities of hydropower, wind, geothermal and solar energies for top countries in the world.



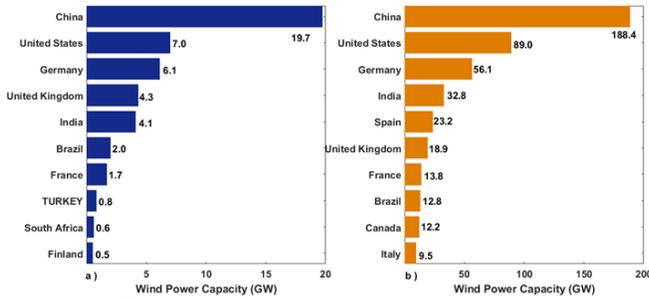


Fig. 3. The new installations (a) and total (b) power capacities of hydropower, wind, geothermal and solar energies for top countries in the world [23].

The present study evaluates the utilization of renewable energy in Turkey’s energy production, especially for solar energy. The aim of the study is to focus on the production and rising of the solar energy in the electricity generation in Turkey. The present policy and legal aspects on of the feed-in-tariff of photovoltaic systems in Turkey is also explained in this research. Furthermore, the effect on solar energy utilization of the incentives and project calls supported by legal arrangements are investigated in the field of solar power plants.

II. STATUS AND PROGRESS OF TURKEY’S PRODUCTION AND CONSUMPTION OF ENERGY

Rapid population growth and industrialization significantly increase energy demands in developing countries. Energy, as a production factor, is one of the main indicators that can be used to project the economic and social development potential of any country. There exists a linear relationship between energy consumption and social development [29]. In 2017, energy consumption in Turkey was 111,782,000 Ton Equivalent Petroleum (TEP). Specifically, within the year 2017, 44,319,000 TEP of natural gas, 13,791,000 TEP of lignite and 24,707,000 TEP of fat coal were consumed in the country. The consumption rate of natural gas was 53,720 million per m³ in 2017, whereas its importation rate was 99% of the natural gas consumed for energy production. The imported natural gas also had an important role in the production of electrical energy as seen in Fig. 4 [30].

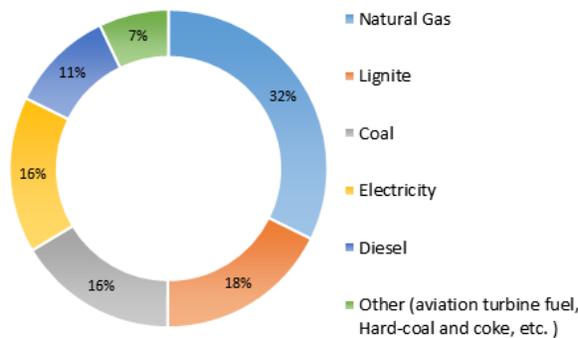


Fig.3. The distribution of the total consumption of the energy by sources in Turkey [27].

According to the 2016 report of the World Energy Council, most of the energy used globally was obtained from primary energy sources. The resources with the largest share in primary energy use were petrol (32.94%), coal (29.20%) and natural gas (23.85%) [31].

Fig. 5 shows the distribution of energy resources used for the generation of electrical energy between 2006 and 2017 in Turkey. The electrical energy generated in 2006 was obtained from natural gas, lignite, coal and petrol products, at rates of 46%, 25%, 18% and 21%, respectively. The use of RESs, other than hydraulic energy, was negligible in 2006. By the year 2017, these proportions reached 35% for natural gas, 19% for hydraulic, 17% for coal, 13% for lignite, 7% for solar 6% for wind and 5% for other sources.

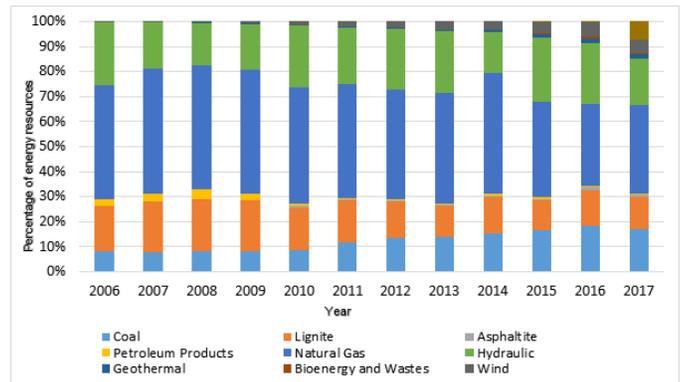


Fig.5. Electricity generation rates according to types of energy resources between 2006-2017 (%) [29].

Table 1 presents the energy generation data of the resources used to generate electrical energy in Turkey between the years 2006 and 2017 (GWh) [30]. As shown in Table I, recently, the production rates of RESs such as wind, solar, geothermal and biomass have been increasing year by year. In particular, the rates of wind and solar energy in energy production have been significantly greater as compared with other energy resources. Electricity generation from solar energy has increased about 20 times reaching to 20,291 GWh in 2017, owing to incentives and investments, while it was 1,043 GWh in the previous year [30].

TABLE I
THE RESOURCES USED IN THE ELECTRICAL ENERGY GENERATION BETWEEN 2006-2017 IN TURKEY (GWh)

<i>Types of Energy Resources (GWh)</i>											
<i>Year</i>	<i>Coal</i>	<i>Lignite</i>	<i>Asphaltite</i>	<i>Petroleum Products¹</i>	<i>Natural Gas²</i>	<i>Hydraulic</i>	<i>Geothermal</i>	<i>Bioenergy and Wastes</i>	<i>Wind</i>	<i>Solar</i>	<i>Total Energy</i>
2017	54,387	40,694	2,395	1,200	110,490	58,218	6,127	2,889	17,904	20,291	297,278
2016	50,829	38,570	2,874	1,926	89,227	67,231	4,819	2,372	15,517	1,043	274,408
2015	43,751	31,336	1,079	2,224	99,219	67,146	3,425	1,758	11,653	194	261,784
2014	38,693	36,615	954	2,145	120,576	40,645	2,364		8,520	17	251,963
2013	32,792	30,262	732	1,739	105,116	59,420	1,364		7,558		240,154
2012	32,475	34,689	850	1,639	104,499	57,865	899		5,861		239,497
2011	26,531	38,870	817	904	104,048	52,339	694		4,724		229,396
2010	18,120	35,942	984	2,180	98,144	51,796	668		2,916		211,208
2009	16,148	39,089	448	4,803	96,095	35,959	436		1,495		194,813
2008	15,858	41,858		7,519	98,685	33,270	162		847		198,418
2007	15,136	38,295		6,527	95,025	36,007			355		191,558
2006	14,217	32,433		4,340	80,691	44,338			127		176,300

¹It is the sum of Petroleum Odor, Fuel Oil, Motor, Gasoline, LPG, Refinery Gas, Aviation Fuel, Gas Oil, Naphtha, Intermediates, Base Oil, White Alcohol, Bitumen, Marine Motor, Marine Fuel and others.

²9155 kcal/Sm³ at 15 °C and 1 atmosphere pressure.

Coal and lignite are the primary local energy resources in Turkey in addition to hydropower and the other RESs. Turkey does not have too many reserves of natural gas and/or other fossil resources. Table II shows the installed power capacity in

Turkey. In 2017, extracted coal and lignite formed about 23% of Turkey's primary energy resources, whereas its renewable energy capacity was about 48% of the total energy in the installed capacity [30].

TABLE II
INSTALLED POWER CAPACITY BETWEEN 2006-2017 IN TURKEY

<i>Types of Energy Resources (MW)</i>											
<i>Year</i>	<i>Coal</i>	<i>Lignite</i>	<i>Asphaltite</i>	<i>Petroleum Products¹</i>	<i>Natural Gas²</i>	<i>Hydraulic</i>	<i>Geothermal</i>	<i>Bioenergy and Wastes</i>	<i>Wind</i>	<i>Solar</i>	<i>Total Power</i>
2017	9,171	9,090	405	986	26,639	27,273	1,063.7	634	6,516	3,421	85,200
2016	7,824	9,270	405	645	25,771	26,681	821	496	5,751	833	78,497
2015	6,690	8,697	135	1,105	24,906	25,868	624	370	4,503	249	73,147
2014	6,398	8,281	135	1,181	25,508	23,643	405		3,630	40	69,520
2013	4,248	8,223	135	1,229	24,579	22,289	311		2,760		64,008
2012	4,248	8,193	135	1,884	20,399	19,609	162		2,261		57,059
2011	4,216	8,199	135	1,778	19,477	17,137	114		1,729		52,911
2010	3,616	8,199	135	2,046	18,175	15,832	94		1,320		49,524
2009	2,256	8,110	135	2,140	16,617	14,553	77		792		44,761
2008	1,986	8,205		2,290	15,055	13,829	30		364		41,817
2007	1,986	8,211		14,576	14,576	13,418			146		40,836
2006	1,986	8,211		2397	14,331	13,086			59		40,565

¹It is the sum of Petroleum Odor, Fuel Oil, Motor, Gasoline, LPG, Refinery Gas, Aviation Fuel, Gas Oil, Naphtha, Intermediates, Base Oil, White Alcohol, Bitumen, Marine Motor, Marine Fuel and others.

²9155 kcal/Sm³ at 15 °C and 1 atmosphere pressure.

The Republic of Turkey Energy Market Regulatory Authority (EMRA) and related organisations classify electrical energy production according to organisations as Transfer of

Operational Right, Build Operate, Build Operate Transfer, Free Production Company, Electricity Generation Corporation (EUAS) and Auto-producers. Fig. 6 shows the distribution of

electricity generation in 2016. Free production companies have the largest share in electric energy production with a rate of 61.48%. The second organization with the largest share is EUAS, with a rate of 26% [32].

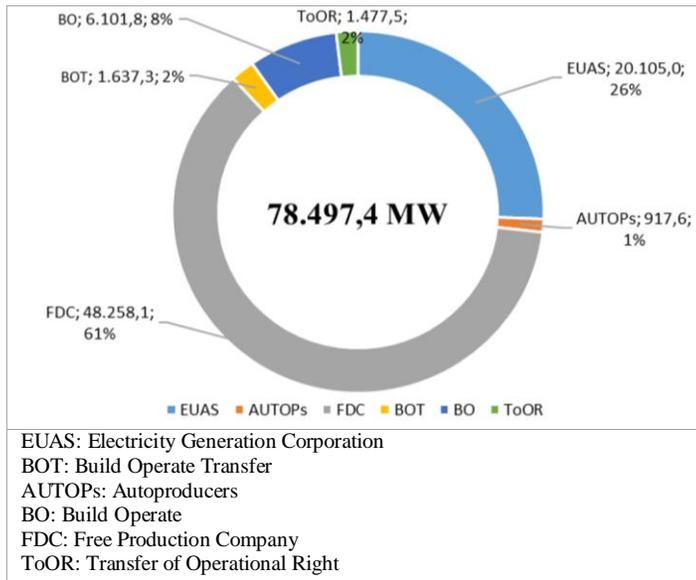


Fig.6. Distribution of electricity generation in 2016 (%) [30].

An examination of the private sector distribution of electrical energy production reveals that the public sector's total share (including current contracts) in the production was 60.3% in 2006, and decreased to 25.6% in 2016. Conversely, the private sector's total share in the production was 39.7% in 2006, and increased to 74.4% in 2016 (See Fig. 6) [33]. As shown in Fig. 7, the public sector's share is decreasing day by day compared to past years. This decrease implies that a large part of the system's newly added electrical power plants is owned by private companies, and/or some of the existing power plants have been handed over to the private sector for operation.

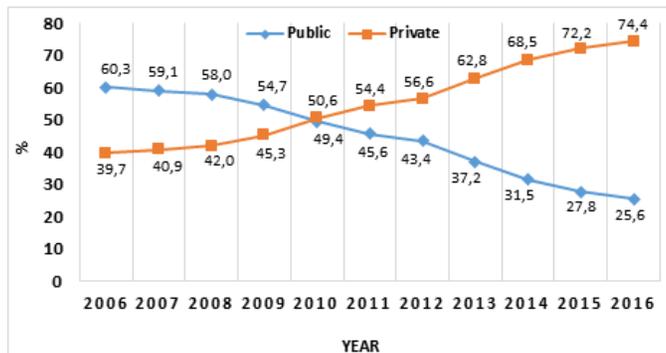


Fig.7. Public-private sectors variation of the production of the electrical energy between 2006-2016 (%) [31].

Fig. 8 shows the rates of domestic and imported resources in Turkey's installed power for the years 2000–2016. The total share of imported resources in the production was 33.7% in 2000 and increased to 43.4% in 2016. The share of domestic resources in 2000 was 66.3% and decreased to 56.6% in 2016 [33].

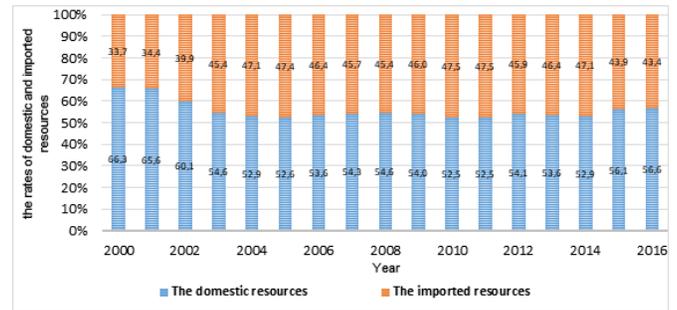


Fig.8. The rates in the use of the domestic and the imported resources in Turkey's installed power between 2000-2016 (%) [31].

Turkey has exhibited an average economic growth rate of 5% in the last 15 years. This growth rate causes rapid increase in energy demands. Yet, the use of natural gas plants, whose installation time and cost are low, has increased. However, in the near future, the use of imported resources is expected to decrease, owing to local investment and incentives that have been provided for the RESs exploitation in recent years.

III. UTILIZATION OF RENEWABLE ENERGY ON TURKEY'S ENERGY PRODUCTION

Following the global trend, significant progress is witnessed in Turkey's renewable energy utilisation. Due to its unique geographical location, Turkey has the opportunity to benefit from all renewable energy sources. Fig. 9 presents Turkey's installed power capacity with respect to a comparison between the primary source of energy and other resources, for 2017. The comparison reveals that the installed power capacity of all resources except liquid fuels has increased between 2006 and 2017. Upon examining RESs, hydraulic energy is also observed to have a greater rate of installed power capacity, followed by wind energy, solar energy and geothermal energy.

Whereas theoretical potential of hydropower resources having the most important place in Turkey's renewable energy potential is 433 TWh/year, assessable and economic potentials are 216 and 140 TWh/year, respectively [30]. Turkey's theoretical hydroelectric potential corresponds to 1% of the world theoretical potential. Besides, its economic hydroelectric potential also corresponds to 16% of Europe's economic hydroelectric potential [30]. In recent years, investments of electricity production from hydraulic energy source has been canalised to small-scale power plants. By the end of 2017, 628 active hydropower cumulative installed capacity of power plants in Turkey was 27,265 MW [30]. This value corresponds to approximately 32% of the total installed capacity of Turkey as seen in Table II.

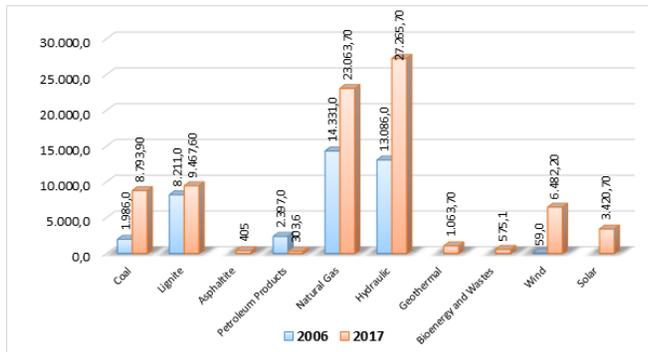


Fig.9. Turkey installed capacity according to the primary energy source for 2006 and 2017 [31].

The Turkish Parliament's regulations have come into force in order to increase the renewable energy utilisation in electricity production, since 2005. For this reason, the installed wind capacity increased more than 100 times to 6,482 MW by end of 2017. The Republic of Turkey Ministry of Energy and Natural Resources (MENR) has set a 2023 wind energy target of up to 20,000 MW. The production of the electrical energy from wind turbines was 15,524 GWh in 2017 [30].

Turkey has significantly increased the utilization of geothermal energy in the past few years. This increase is substantially due to the geothermal law with favouring regulations, feed-in tariffs and a target of 1 GW of geothermal power capacity to be installed by 2023. It is reported that 225 new geothermal fields have been discovered in year 2015. A total of 10 new power units of 159 MW have been completed and total installed power capacity has risen to 1,063 MW [30].

In addition to other renewable energy resources, Turkey invests on bioenergy conversion technologies, as well. There are 122 Bioenergy Power Plants with 575.1 MW installed power, which correspond to 0.7% of installed power capacity of Turkey by end of 2017. Electricity production using bioenergy and wastes was 1,758 GWh, corresponding to 0.95% of total electricity production of Turkey in 2017 [30].

Significant improvements are observed in Turkey's renewable energy utilisation, especially in the past 15 years. Incentives, investments and legal regulations paved the way for this progress. Consequently, the ratio of renewable energy in the total generated energy is reached to about 1/2 in 2017. This figure was 1/4 in 2006.

IV. PRESENT AND FUTURE UTILIZATION OF SOLAR ENERGY IN TURKEY

Solar energy as an RES has greater potential in Turkey than in other European countries. However, this potential has remained largely unexploited. Nevertheless, in recent years, incentives have been granted and project calls have been opened to investors aiming to increase the use of RESs, especially solar energy. Consequently, solar energy is expected to significantly contribute to Turkey's installed power in the near future.

A. Solar Energy Potential in Turkey

Turkey has a good geographical location, and therefore high solar energy potential. Fig.10 shows the distribution of Turkey's average global horizontal irradiance between 1994 and 2015 [34]. Using the data obtained from MENR, it can be said that Turkey has an average annual total sunlight duration of 2,737 hours (7.5 h per day on average), an average total radiation intensity of 1,527 kWh/m²-year (daily total of 4.2 kWh/m²-day). In terms of months, maximum solar energy can be produced in June and July, and minimum solar energy can be produced in December and January [35]. The region with the highest solar radiation in Turkey is the Southeastern Anatolia Region, as seen in Fig. 11, followed by the Mediterranean Region. The region with the lowest potential for solar radiation is the Black Sea region with a total radiation intensity of about 1200 kWh/m²-year [35]. However, the total radiation intensity in this region is quite higher than that in most European countries.

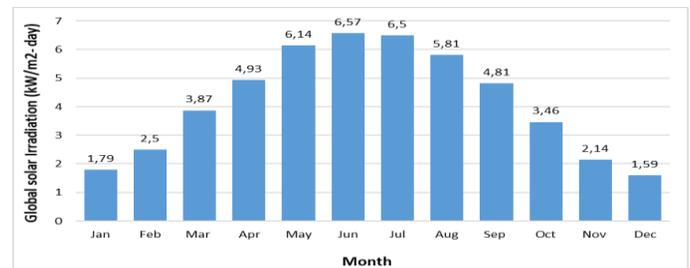


Fig. 10. Average global solar irradiation per m²-day in Turkey.

Solar power plant technologies can be divided into two main groups, widely varying in terms of method, material and technological level:

- Thermal Solar Technologies and Concentrating Solar Power: Heat derived from solar energy can be used directly or indirectly for electricity generation in these systems.
- Photovoltaic Systems: These systems are based on semiconductor-based devices that directly convert solar radiation to electricity.

Turkey takes an important place in the world in the use of solar thermal capacity. In 2017, an estimated 35 GW of new solar thermal energy plant was commissioned in the world. China is the leading country in the new installation, followed by Turkey, India, Brazil and the United States [23]. The total installed solar collector area in Turkey is estimated to be approximately 20 km². About 50% of the produced planar collectors and all vacuum tube collectors are used in the country. Approximately 823,000 TEP of heat energy was produced using solar collectors in 2017. Residential use of the produced heat energy was 536,000 TEP and industrial use was 287,000 TEP [35].

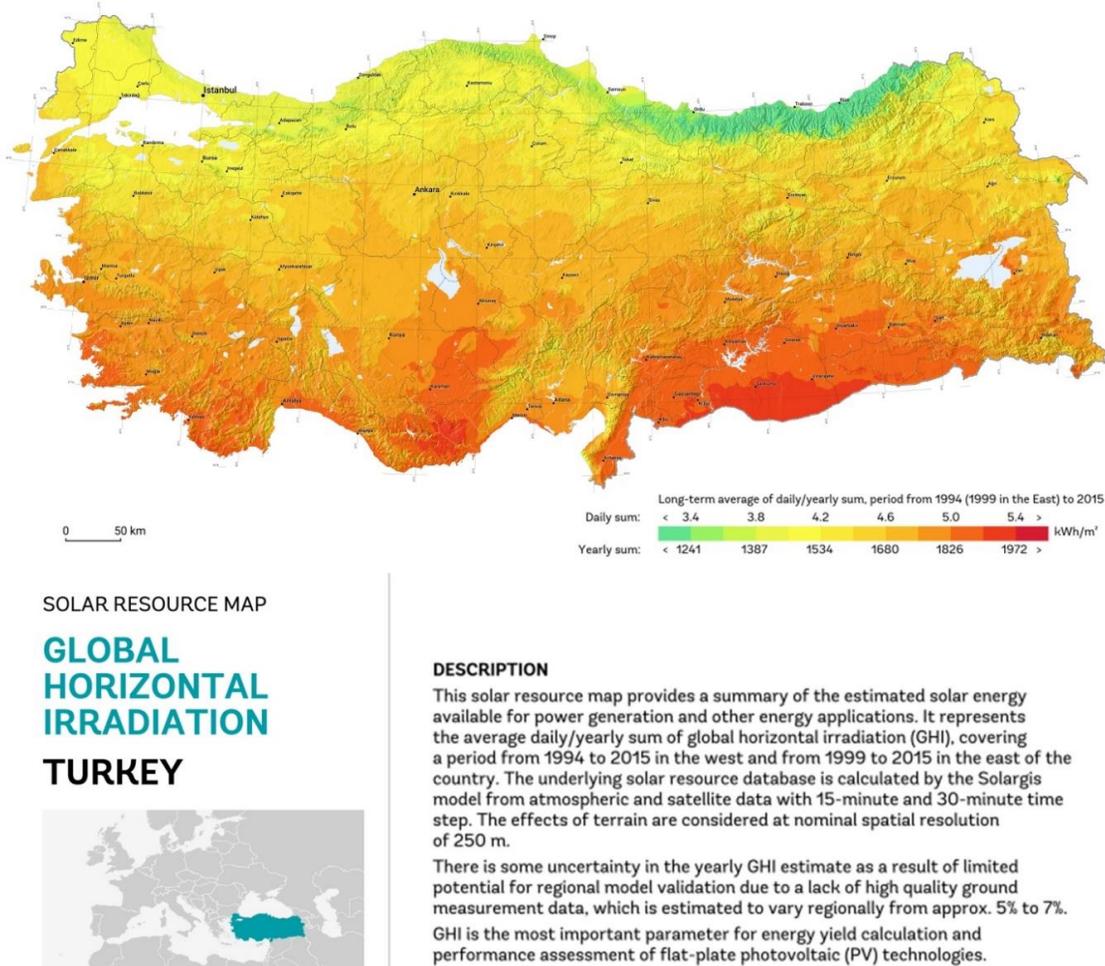


Fig.11. Turkey’s global solar atlas source: World Bank Group [32].

B. Present Policy and Legal Aspects of the Feed-in-Tariff of Photovoltaic Systems in Turkey

In recent years, awareness has grown parallel to technological developments leading to incentives and investments in photovoltaic systems in Turkey. Therefore, many domestic and foreign companies in the energy sector are interested in policies on investment issues concerning solar power plants in Turkey.

In recent years, the investment capacity for solar power systems is increasing on a daily basis with announced incentive prices and 10-year purchase guarantee. To support entrepreneurs who wish to invest in this sector in Turkey, Renewable Energy Law No. 5346, Electricity Market Law No. 6446 and Unlicensed Electricity Production Regulation have been modified. With current arrangements, the value of 1 kWh of electricity produced by solar power plants operating until 31 December, 2020 for all real and legal persons will be 13.3 cent/kWh. However, this price increases up to 20 cent/kWh depending on the domestic devices used in the construction of the power plant. If all devices are domestic, the price is 20 cent/kWh. Solar power plant license applications for 600 MW capacity were received by EMRA in 38 cities covering 27 regions between 10 and 14 June, 2013. Furthermore, 496 solar power plant license applications were received for 7,930 MW capacity. Those eligible for investments are entitled to a

license in these applications. Fig. 12 shows the participation trend in the competition for electricity generation from solar energy.



Fig.10. Map showing incentive zones for the installation of solar power plants covering 27 districts 38 cities for 600MW capacity received from EMRA [30].

Up to 1 MW of electricity can be produced without a license for solar power plants within the scope of the Regulation on Unlicensed Electricity Generation in Electricity Market, dated 2 October, 2013 and numbered 28783. Moreover, investors are encouraged by the government for roof and solar power plant system applications. The capacity entering the operation, which was 359.04 MWe at the end of 2015, in the production of unlicensed electricity has increased about three times to 1,048.21 MWe as of the end of 2016.

Herein, the largest share belongs to solar power plants with 89.60%, followed by natural gas at 4.95% and biomass at 3.47% [32].

C. Capacity of Photovoltaic Systems for Energy Production in Turkey

In terms of solar energy production, Turkey has a better position as compared with many developed and developing countries owing to its geographical location (a sunny belt between 36°N and 42°N). In an age when energy demands are increasing parallel to the development of new technology, the use of renewable energy-based systems instead of traditional ones is expected to diminish the country's dependence on foreign energy sources. For this reason, incentives and project calls, supported by the corresponding legal arrangements, in the field of solar power plant applications have encouraged many domestic and foreign investors to invest in this field in Turkey.

Turkey's cumulative installed capacity of solar power plants (40 MW), including licensed and unlicensed plants, formed 0.06% of the total installed power plant capacity in 2014, according to the data provided by the MENR. As of the end of 2017, Turkey's cumulative installed power capacity (3421 MW) increased to 4.01% of the total installed power capacity with completed licensed and unlicensed plants supported via incentives and projects calls in the field of solar power plant applications. The installed power capacity in 2017 increased by about 85 times compared to the year 2014.

The production-capacity-projection-model study of MENR predicted the total installed power by type of energy source, in accordance with macroeconomic targets, between 2016 and 2020 (Turkish Electricity Transmission Company). The capacity of solar power plants (in power plants, public facilities under construction, Licensed-Associate/Licensed/Unlicensed/Solar Production Facilities within the scope of the (Renewable Energy Resource Area) RERA Project) for the year 2016, was estimated at 605.3 MW according to this report [36]. The installed capacity at the end of 2016 was 833 MW. Therefore, the existing installed power was estimated to be 37.62% higher than expected. The installed capacity for 2020 was determined to be 4,084.7 MW [36]. This capacity already has increased about 4 times to 3,421 MW in 2017, owing to incentives and investments when compared to the previous year. Solar power plant capacity is expected to reach well above the projected value in 2020.

1,500 MW solar photovoltaic power plants are expected to be installed over an area of 27.2 km² declared as the RERA in the Karapinar district of the Konya province (Central Anatolia Region). Installation will be performed in two stages, providing 1,000 MW and 500 MW of energy (Karapinar Energy Specialised Industrial Zone Part 1), as shown in Fig. 13.

The construction of an 1,800 MW solar photovoltaic power plant is planned for the tender's second part concerning the 32.4 km² area declared as Karapinar RERA (Karapinar Energy Specialised Industrial Zone Part 2). Part 1 of the tender by the Energy and Natural Resources Ministry was realised through the participation of four consortiums in 2017. The Kalyoncu–Hanwha Group submitted the lowest bid of 6.99 cent/kWh in the tender, which was completed in 18 rounds.

Under this bid, a factory installation with a minimum production capacity of 500 MW PV modules per year will be implemented in Turkey.

In the Official Gazette dated Sep. 29, 2018 and numbered 30550, new RERA areas were identified. These areas, Niğde/Bor, Şanlıurfa/Viransehir and covering the boundaries of three district Hatay/Erzin Adana/Toprakkale and Osmaniye/Ceyhan are the three RERA areas.

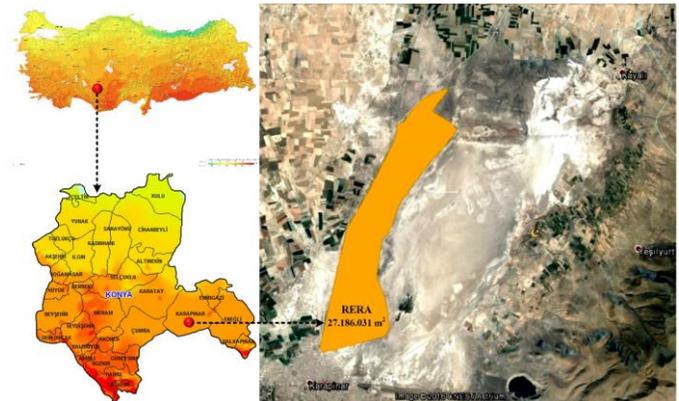


Fig. 13. Renewable Energy Resource Area (RERA): Karapinar Energy Specialized Industrial Zone Part 1, 27.186.031m², Konya, Turkey [33].

In this study, more than 700 solar photovoltaic power plants in Turkey are identified and mapped. For the mapping, the power plants are divided into four categories: active, in the construction phase, in the office phase and planned power plants. By the end of 2017, active solar photovoltaic cumulative installed capacity of power plants in Turkey was 3,421 MW, exhibiting an increase of 4.11 times over the previous year. Solar power installed capacity at the regional scale is mostly focused in the south of the Central Anatolian region, followed by the Mediterranean and Aegean regions as seen in Fig. 14. When evaluated on a provincial basis, installed power capacity is observed to be the highest in Konya, followed by Kayseri and Ankara.

Although solar radiation is the most intense in the Southeastern Anatolia Region, there are more solar photovoltaic power plants in Central Anatolia and in the Mediterranean Region. The criteria for solar power plant installation depend on land choice (no-ghost effect, south-facing and flat areas with less than 8% slope, geological and hydrological factors, not in agricultural fields, etc.), environmental climatic conditions (temperature, humidity, wind, snow, etc.), geopolitical risk factors (such as terrorism) and proximity to transformer centers. Thus, land structure and geopolitical risk factors play important roles in the less-than-expected solar photovoltaic power plant capacity of the Southeastern Anatolia Region. Furthermore, more solar photovoltaic power plants exist in the construction stage in the Central Anatolia, Mediterranean and Aegean regions.

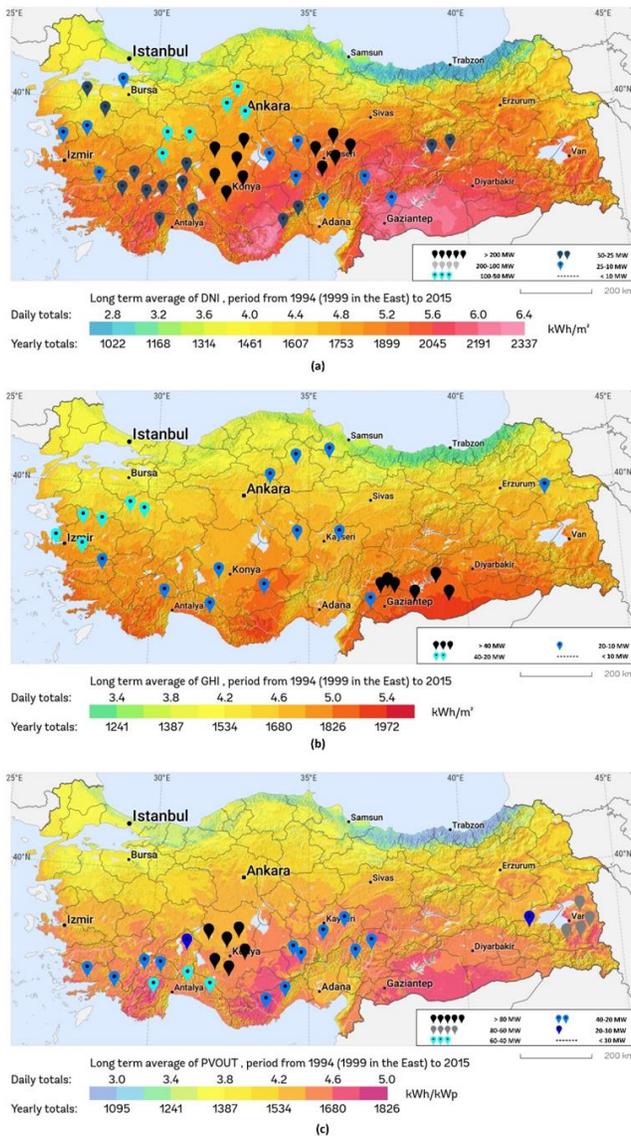


Fig.14. Map showing of Solar Photovoltaic Power Plant Capacity (MW) as active (a), construction phase (b), office phase (c).

V. CONCLUSION

Rapid population growth and industrialisation in Turkey have resulted in a rapidly increasing energy demand. This energy demand reflects the increased potential for economic and socio-cultural development. Different types of energy resources are used to meet the energy demand of Turkey. In this regard, the most effective and reliable sources of energy are RESs that do not produce harmful carbon emissions. Due to the incentives provided and the investments made in the last few years, the share of renewable energy in the total production of energy increased from 1/4 in 2006 to about 1/2 in 2017. Thus, the capacity utilisation rate of RESs has increased extremely rapidly. By the end of 2017, the installed power capacity of solar photovoltaic power plants increased by 4.11 times compared to the previous year, reaching 3,421 MW. This value corresponds to 4% of the total installed power. The installed capacity for 2020 was determined to be 4,084.7 MW. Because this capacity already has increased to

3,421 MW in 2017, solar power plant capacity is expected to reach well above the projected value in 2020 owing to incentives and investments. At a regional scale, solar power installed capacity is mostly located in the south of the Central Anatolian region, followed by the Mediterranean and Aegean regions. When evaluated on a provincial basis, the installed capacity is mostly focused in Konya, followed by Kayseri and Ankara. Although the Southeastern Anatolia Region has the highest solar radiation levels in Turkey, solar photovoltaic power plants density is highest in Central Anatolia and the Mediterranean region. The terrain selection criteria for solar power plants include land structure, climate, geopolitical risk factors and proximity to substation centres. Consequently, land structure and especially geopolitical risk factors play a crucial role in the observed lack of solar photovoltaic power plants capacity in Southeastern Anatolia. As a result, supported by the required legal arrangements made by the concerned public and other related institutions, and owing to project calls and incentives for solar power plants, many domestic and foreign investors have been encouraged to work in Turkey. In this context, solar energy power capacity is expected to significantly contribute to the total installed power capacity over the next few years.

REFERENCES

- [1] Pérez-Lombard L, Ortiz J, Pout C. A review on buildings energy consumption information, *Energy and Buildings*, 2008, 40(3):394-398.
- [2] Kalogirou SA. *Solar energy engineering: process and systems*, vol. 1st, California: Academic Press, 2009.
- [3] Tukenmez M., Demireli E. Renewable energy policy in Turkey with the new legal regulations, *Renewable Energy*, 2012,39:1-9.
- [4] Guo S, Liu Q, Sun J, Jin H. A review on the utilization of hybrid renewable energy, *Renewable and Sustainable Energy Reviews*, 2018, 91:1121-1147.
- [5] Esen, V, Neşe, S, Sağlam, Ş, Oral, B, "The Training of Renewable Energy Systems Undergraduate Studies". *Balkan Journal of Electrical and Computer Engineering* 5 (2016): 26-29.
- [6] Sahin, F. E., Musa Yilmaz. "High concentration photovoltaics (HCPV) with diffractive secondary optical elements." *Photonics*. Vol. 6. No. 2. Multidisciplinary Digital Publishing Institute, 2019.
- [7] Montoya FG, Aguilera MJ, Manzano-Agugliaro F. Renewable energy production in Spain: A review, *Renewable and Sustainable Energy Reviews*, 2014, 33:509-531.
- [8] Yılmaz, M, "Real Measure of a Transmission Line Data with Load Forecast Model for The Future". *Balkan Journal of Electrical and Computer Engineering* 6 (2018): 141-145
- [9] Tur, M, "The Impact of Emerging Renewable Energy on Capacity Mechanisms in Power Systems and Expert Opinion". *Balkan Journal of Electrical and Computer Engineering* 7 (2019): 319-325
- [10] Salam M, Yazdani M, Rahman Q, Nurul D, Mei SF, Hasan S. Investigation of wind energy potentials in Brunei Darussalam, *Front. Energy*, 2018, 125:1-11.
- [11] Dirks L, Dirks G, Wu J. Evolving perspectives on biofuels in the United States, *Front. Energy*, 2012, 6:379-393.
- [12] Yuksel I. Energy production and sustainable energy policies in Turkey, *Renewable Energy*, 2010, 35: 1469-1476.
- [13] World Bank, "World Bank, World Development Indicators (WDI)," *Our World in Data*, 2014. <https://ourworldindata.org/grapher/global-electricity-production-by-source,2018-1-24>.
- [14] Goetzberger A, Hoffmann V. *Photovoltaic Solar Energy Generation*, Germany: Springer-Verlag Berlin Heidelberg, 2005.
- [15] UNFCCC, "Climate change: Climate change: Impacts, vulnerabilities and adaptation in developing countries" United Nations Framework Convention on Climate Change, Bonn, Germany, 2007.

- [16] Owusu P A, Asumadu-Sarkodie S. A review of renewable energy sources, sustainability issues and climate change mitigation, *Cogent Engineering*, 2016, 3(1167990):1-14.
- [17] Goel A, Bhatt R. Causes and Consequences of Global Warming, *Int J LifeSc Bt & Pharm Res*, 2012, 1(1):27-31.
- [18] IEA, "CO2 Emissions from Fuel Combustion, IEA, 2017," INTERNATIONAL ENERGY AGENCY, 2017. <https://www.iea.org/publications/freepublications/publication/co2-emissions-from-fuel-combustion---2017-edition---overview.html>, 2017-1-24.
- [19] Energy Information Administration (EIA), "In: International Energy Outlook 2009 (IEO2009)," Official energy statistics from the U.S., 2009.
- [20] Turkish Statistical Institute (TUIK), "Energy statistics. Electricity production and distribution." Republic of Turkey: Prime Ministry, Ankara, 2016.
- [21] Nguyen K Q. Alternatives to grid extension for rural electrification: Decentralized renewable energy technologies in Vietnam, *Energy Policy*, 2007,35(4):2579–2589.
- [22] Mekhilef S, Saidur R, Safari A. A review on solar energy use in industries, *Renew Sustain Energy Rev*, 2011,15(4):1777–1790.
- [23] RENEWABLES 2018 GLOBAL STATUS REPORT, Renewable Energy Policy Network for the 21st Century, 2018.
- [24] Hosenuzzaman M, Rahim N, Selvaraj J, Hasanuzzaman M, Malek A, Nahar A. Global prospects, progress, policies, and environmental impact of solar photovoltaic power generation, *Renewable and Sustainable Energy Reviews*, 2015, 41:284-297.
- [25] Solangi K, Islam M, Saidur R, Rahim N, Fayaz H. A review on global solar energy policy, *Ren Sustain Energy Rev*, 2011,15(4):2149–2163.
- [26] Ahmed F, Al Amin A, Hasanuzzaman M, Saidur R. Alternative energy resources in Bangladesh and future prospect. *Renew Sustain Energy Rev*, 2013, 25:698–707.
- [27] Lynn PA. *Electricity from sunlight: an introduction to photovoltaics*. West Sussex: John Wiley & Sons Ltd, 2010.
- [28] Kaya D. Renewable energy policies in Turkey, *Renewable and Sustainable Energy Reviews*, 2006,10: 152–163.
- [29] Koç E, Şenel M C. Dünnyada ve Türkiye’de Enerji Durumu -The State of Energy in World and Turkey - General Evaluation, *Mühendis ve Makina*, 2013;54(639):32-44.
- [30] Ministry of Energy and Natural Resources (MENR), "Electricity generation-distribution statistics according to types of energy resources," Dec. 2017. <http://www.eigm.gov.tr/tr-TR/Denge-Tablolari/Denge-Tablolari>. 2018-2-26.
- [31] WORLD ENERGY COUNCIL, "World Energy Resources | 2016," Copyright © 2016 World Energy Council, London EC3V 3NH, United Kingdom, 2016.
- [32] Energy Market Regulatory Authority (EMRA), "Sector reports," <http://www.epdk.org.tr/TR/Dokumanlar/Elektrik/YayinlarRaporlar/ElektrikPiyasasiGelisimRaporu>, 2016, 2018-2-26.
- [33] Turkish Electricity Transmission Company (TEIAS), "Turkey's Electricity Generation-Transmission Statistics," Dec. 2016. <https://www.teias.gov.tr/turkiye-elektrik-uretim-iletim-istatistikleri>. 2018-02-27.
- [34] World Bank Group, "Global Solar Atlas, Turkey," The World Bank and International Finance Corporation (IFC), 30 01 2017. <http://globalsolaratlas.info/downloads/turkey.2018-02-06>.
- [35] Ministry of Energy and Natural Resources (MENR), "Solar energy capacity in Turkey," Dec. 2016. <http://www.enerji.gov.tr/tr-TR/Sayfalar/Gunes>. 2018-02-27.
- [36] Turkish Electricity Transmission Company (TEIAS), "Turkey's electricity generation capacity for 5-year projection (2016-2020)," Ankara, 2015.

BIOGRAPHIES



FARUK ERKEN- He received the BS degree from the Department of Electrical and Electronics Engineering, Firat University, the MSc degree from Institute of Science and Technology, Dicle University, the PhD degree from Institute of Science and Technology, Firat University respectively, from Turkey. He is currently an Assistant

Professor in the Electrical and Electronics Engineering Department of Kastamonu University His research interests include electrical machines and control, power converters and renewable energy.

Automated Fake Access Point Attack Detection and Prevention System with IoT Devices

I. F. KILINÇER, F. ERTAM and A. ŞENGÜR

Abstract— Wireless access points (APs), which allow many devices to be easily connected to the Internet, are widely used today because they offer the easiest way to connect to the Internet. With the development of the Internet of Things (IoT), WiFi networks are widely used in our homes, workplaces, social areas, campus areas. With the increase of WiFi networks, attacks on these networks are constantly increasing. In this study, an IoT-based approach to detect and prevent Fake Access point attacks frequently seen in WiFi networks is proposed. A Single Board Computer (SBC) and a wireless antenna in the "Soft AP" feature are used for operation. Fake APs were detected by air scanning. In the first phase of the study, fake Access point broadcasts have been created which can create security weakness. In order to determine the fake Access points created in the second stage, SBC and wifi module were used to scan air. In the final stage, the mac address of the fake AP has been assigned to an unauthorized Virtual Local Area Network (vLAN) on the network to prevent detected fake AP broadcasts. The possible attack methods for the study were implemented and it was revealed that the proposed approach prevented the attack successfully in all scenarios. The study is seen as an effective, developed and economically useful IoT application for network administrators to prevent the attack using fake Access point.

Index Terms— Attack detection; Attack prevention; DoS attack; Network security, Fake access point, IoT

I. INTRODUCTION

CONSIDERING THE increase in the number of devices connected to the Internet and the lack of the possibility of connecting to the Internet via cable on all of these devices, the use of wireless access networks such as WiFi has emerged as a practical solution [1]. In particular, Internet of Things (IoT),

İLHAN FIRAT KILINÇER, is with Department of Informatics of Firat University, Elazig, Turkey, (e-mail: ifkilincer@firat.edu.tr).

 <https://orcid.org/0000-0001-8090-4998>

FATİH ERTAM, is with Department of Digital Forensics of Firat University, Elazig, Turkey, (e-mail: fatih.ertam@firat.edu.tr).

 <https://orcid.org/0000-0002-2306-6008>

ABDÜLKADİR ŞENGÜR, is with Department of Electrical and Electronical Engineering of Firat University, Elazig, Turkey, (e-mail: asengur@firat.edu.tr).

 <https://orcid.org/0000-0002-2306-6008>

Manuscript received October 16, 2019; accepted Nov 13, 2019.

DOI: [10.17694/bajece.634104](https://doi.org/10.17694/bajece.634104)

the way in which the interaction of human and technology changes a little more. Wireless WiFi networks, one of the most important areas of IoT; one or more wireless devices communicating through electromagnetic waves in the same environment. Wireless WiFi technology is the IEEE 802.11 standard and is in the microwave and radar category in the frequency spectrum. Fig. 1 shows the frequency spectrum in which WiFi networks operate [2].

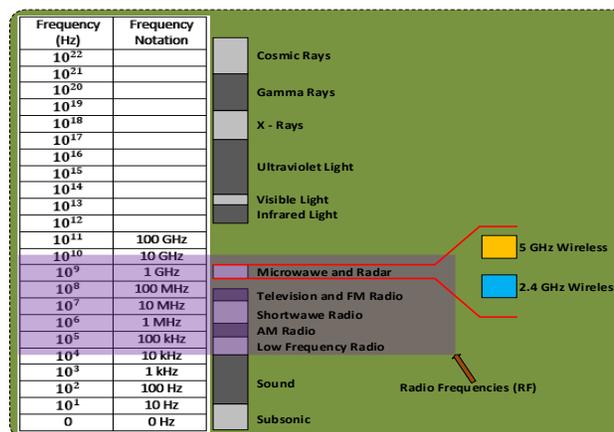


Fig.1. Frequency Spectrum

Fig. 1 shows that the WLAN technology uses the 2.4 GHz and 5 GHz bands. However, the entire band was not used when using these bands. For the 2.4 GHz band; frequencies between 2.400 GHz - 2.4835 GHz 5.7 are used, the frequencies between 5.4 5.150 GHz - 5.250 GHz, 5.250 GHz - 5.350 GHz, 5.470 GHz - 5.725 GHz, 5.725 GHz to 5.825 GHz are used for the 5 GHz band.

WiFi technology offers many important advantages over wired connection because of its ability to make roaming and its comfortable use. However, it contains many security weaknesses. Various security mechanisms such as WiredEquivalentPrivacy (WEP), WiFiProtected Access (WPA), WPA2 have been developed to minimize these security vulnerabilities [3-4].

Heartfield et al. [5], classified the security vulnerabilities that occurred in smart homes, which is an IoT application, taxonomically. Not only the attack vectors were emphasized in the study, but also the potential impact of the attack on the system.

WEP is an encryption algorithm developed to provide security in wireless networks. WEP first started with 64-bit encryption and enabled up to 256-bit encryption. However, the

most commonly used type of encryption is the 128-bit encryption method [6]. In spite of many corrections and increased password dimensions in WEP encryption algorithm, it can be easily broken by many open source applications such as aircrack. Therefore, it is not used much today. The WEP encryption method performs the verification in four steps as in Fig. 2. The WEP encryption method performs the verification in four steps as in Fig. 2. Accordingly, the first step goes from the wireless PC to the connection request to the AP. In the second step, the AP creates a random text (64 bit, 128 bit) according to the encryption method and sends the connection to the requesting device. In the third step, the PC encrypts the text from the AP with the WEP password in itself and sends the encrypted text to the AP. In the last step, if the AP acknowledges the accuracy of the encrypted information that is sent to it, authentication takes place and the user connects to the AP [7-8].

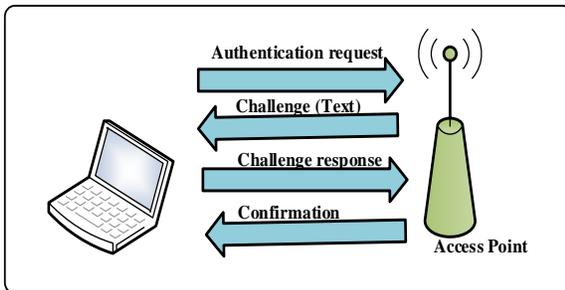


Fig.2. WEP Authentication

The WEP validation method contains many features. Therefore, user data can be easily sniffed by the attacker. The WiFi alliance association then introduced the IEEE802.11i standard to the industry to prevent these vulnerabilities. WPA and WPA2 encryption methods were developed in 2003 with this standard. WPA / WPA2 standards use the Advanced Encryption Standard (AES) based Temporal Key Integrity Protocol (TKIP) and Counter Modewith Cipher Block Chaining Message Authentication Code Protocol (CCMP) encryption methods for secure authentication [6]. In Fig. 3, the classes of the IEEE 802.11i standard are summarized [10].

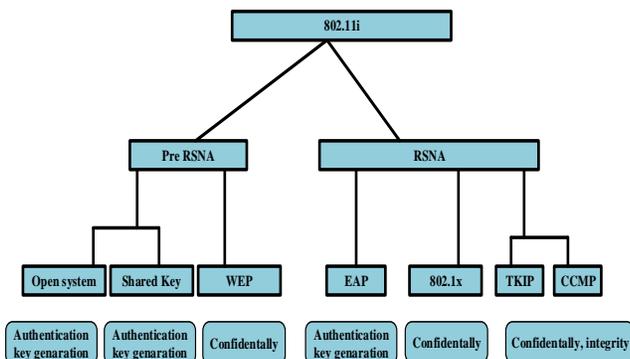


Fig.2. Classes of the 802.11i standard

There are many studies in literature to prevent WEP and WPA cryptography. L YongLei et al. In order to overcome

processor constraints in bruteforce attacks, they proposed distributed multi-core CPU and GPU parallel cracking method (DMCG) method. In this method, WPA has used colored Petri nets to verify the 4-way handshake protocol used in the encryption method. For the DMCG, the generated PSK wordlist is given to each computer running in distributed architecture. GPU processors have also been used to improve performance [11]. V.Kumkar et al. showed how to attack WEP, WPA, WPA2 encrypted networks. WEP and WPA2 networks were attacked and the security vulnerabilities were investigated in real time. The Aircrack-ng framework was used to organize the attack. In the attack, WEP-encrypted networks were seen to be hacked much easier than WPA-encrypted networks [3]. S. Gold, in his study showed that WEP and WPA passwords can be easily broken using tools such as aircrack-ng and aireplay [12].

Wang et al. by analyzing the deficiencies in WEP and WPA encryption methods, they developed d-WEP and d-WPA-PSK algorithms. According to the developed d-WEP algorithm, the decision is made to determine whether the AP is under attack by looking at the number of ARP requests. On the other hand, in order to prevent d-WPA-PSK attacks, a mechanism has been designed in which the PSK is changed regularly [13].

Apart from the WEP and WPA attacks, one of the methods frequently used by the attackers is creating a fake Access point so that users can view their traffic without making them feel. Attacker fake can do an Access point attack in two ways. In the first method, it can broadcast the name “freeWiFi” which will attract the attention of the users. This allows users to connect to it. In the second method, users can connect to it by making a stronger fake broadcast with the same name as the WiFi network in the location. The attacker can follow two options when making the second method.

Option 1:

Catching a 4-way Handshake: The attacker expects users to be authenticate at first, in a network where users are busy. The attacker then tries to solve the network key by capturing the handshake. 4 way handshake is the IEEE-802.11i standard for secure authentication in wireless networks. Pre-sharedkey (PSK) or 802.1X authentication methods use 4 way handshake encryption. During 4 way handshake, there are 4 messages between client and Access point. In the first step, the AP sends a randomly generated Anonce (Authenticatednumberonce) frame to the client with the message 1. In the second step, after obtaining the Anonceframe, Snonce (Supplicantnumberonce) and PTK are generated and transmitted to the AP together with the MIC as message 2. After receiving the AP Snonceframe in the third step, it generates its own GTK and transmits it to the client as message 3 with its MIC. In the last step, the client sends the AP to acknowledge message 4 and the authentication is complete. The algorithm of 4 way handshake is given in Fig. 4 [14-18].

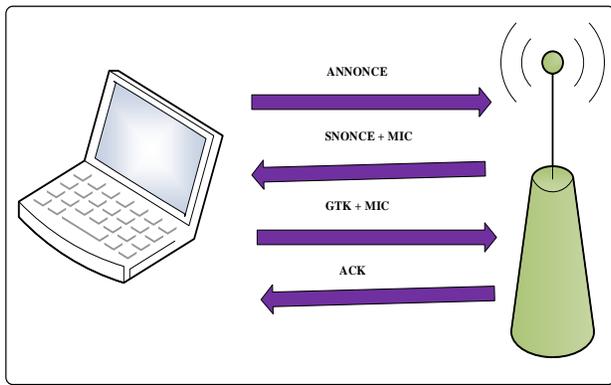


Fig.4. WPA/WPA2 4 way handshake

During the 4-way handshake, GroupTemporalKey (GTK), PairwiseTemporalKey (PTK), Group Master Key (GMK), Pairwise Master Key (PMK), Master SessionKey (MSK), encryption algorithms are used. During authentication, which is the first stage of encryption, the MSK key is generated. In the second stage, the PMK and GMK second level switches produced from the MSK are used to construct the PTK and GTK. The hierarchy of the encryption methods mentioned in Fig. 5 is given.

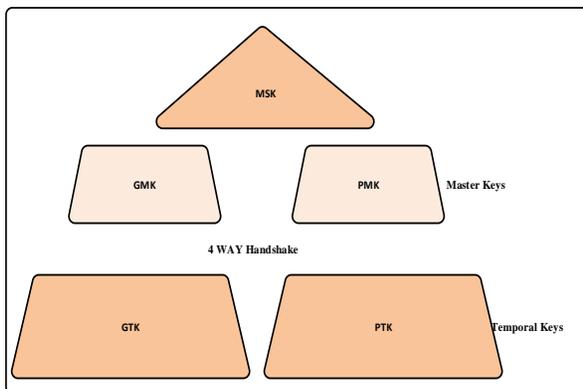


Fig.5. Key Hieararchies

These encryption methods can be briefly summarized as follows [19-20].

- PTK is used to encrypt all unicast traffic between user and Access Point.
- GTK is used to encrypt all broadcasts between Access point and multiple users.
- GMK is used to build GTK.
- The PMK is the key generated from the MSK. The PMK is not usually transmitted over the network, this component is not shared and thus increases transaction security.
- The MSK is the first key derived during 802.1X or PSK authentication.

Option 2:

Deauthentication attack: Deauthentication is a Layer 2 DoS attack specific to the 802.11 standard. The attacker will allow the user to disconnect from the network by organizing a deauthentication attack on the user connected to the AP without waiting for the 4-way handshake. This results in 4-

way handshake when the user wants to connect to the network again. At this stage, the attacker catches the 4-way handshake [14-15,21-22].

Fake AP attacks are frequently used to reach user data. In these types of attacks, the attacker makes a broadcast similar to the SSID in the environment and allows the user to connect to this SSID. There are many studies to prevent fake AP attacks in the literature. Take Kuo-FongKao et al. In order to prevent fake AP attacks, they developed an algorithm by looking at the serial numbers, timestamp and range of beacon messages. In their study, they saw that the attackers could change the serial numbers, time stamps and signal intervals of fake APs. However, they suggested that the method they proposed was successful in detecting these attacks [23].

Mohan K Chirumamilla et al. In their study, they developed an agent-based intrusion detection system to find unauthorized APs. Developed agents maintain a list of registered APs. When an AP is added or removed to any agent, the administrator notifies the MAC address of all AP agents. Thus each agent will have a current AP list on it. Another task of agents is to scan for a fake APs. After each scan, he compares the AP mac addresses he sees with the AP mac addresses in his list. If the AP is not in the current list, the administrator will be notified by SMS. The agent developed in the study has each agent, a wireless interface and two network interface cards. The wireless interface card was used to scan and detect fake APs, while two other Ethernet cards were used to connect the agent to the spine [24]. SomayehNikbakhsh et al. in their study, they proposed a method to prevent users from connecting to fake APs. In the proposed method, the route and gateway information of the package were compared to determine whether an AP is fake. With this method, Man in theMiddle (MiTM) attacks are easily detected without the need for a network administrator [25].

In this study, Raspberry PI 3 was chosen as SBC device to prevent and detect fake AP attacks. Kali Linux was installed on the device, and one of the USB ports was installed on the WiFi antenna.

The detailed contributions are as follows.

- It is very easy to integrate into the existing network, especially in large networks that are used in corporate networks. In this way, a direct ready-to-use system is provided for network administrators and system administrators without any configuration problems.
- To capture fake access point attacks on traditional WiFi networks, there is often a need for an access point that needs to work in monitor mode. Therefore, extra budgets are required to prevent such an attack. The proposed method will work effectively on intrusion detection and prevention without incurring major costs.
- The proposed method will automatically notify the system administrator by means of integrated tools such as SMS and mail when the attack is detected.

II. PROPOSED METHOD

Today, hackers are widely using openings in wireless networks to regulate their attacks. In this context, one of the most common types of attacks in wireless networks is Fake Access point attacks. In this study, a new method has been proposed to prevent fake access point attacks. The network diagram of the proposed method is given in Fig. 6.

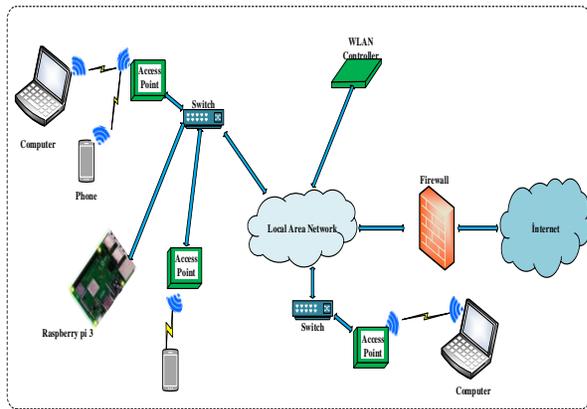


Fig.6. Network diagram of the proposed method

The network diagram given in Fig. 6 is the architecture of a general corporate network. Connected to edge switches in traditional network architectures, the APs transmit incoming traffic to the Wireless controller installed at the center of the network through the LWAPP tunnel.

In this study, a cheaper and effective method has been proposed to prevent fake Access points. Accordingly, instead of an additional Access point, the Raspberry pi device is positioned in the desired area to be monitored and the Wi-Fi broadcasts in the environment are monitored by operating in Raspberry pi monitor mode. Raspberry PI 3 device was used in this study. Kali Linux operating system was installed on 16 GB Class 10 Micro SD card on Raspberry PI 3 device. In this study, ODROID module is installed on the Raspberry PI 3 device for aircan. Raspberry PI and its connections are given Fig.7.

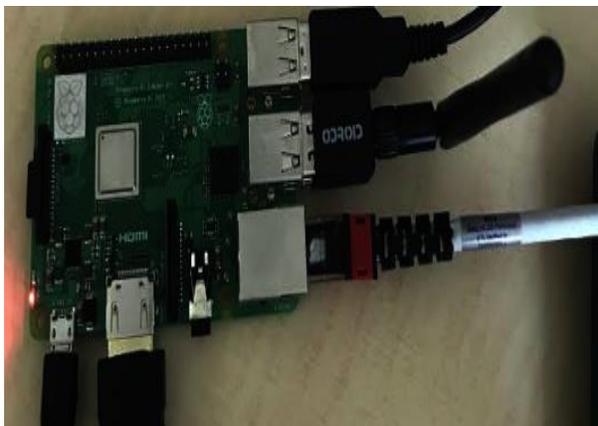


Fig.7. Raspberry PI 3 SBC and connection modules

After the installation of the necessary tools, the algorithm in Fig. 8 was used to find the attacker broadcasting the fake Access point.

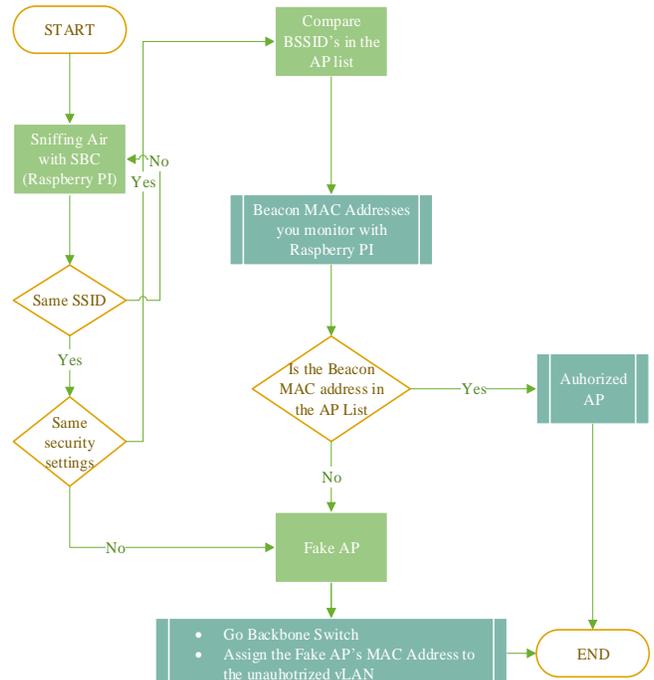


Fig.8. Algorithm of the proposed method

The steps of the flowchart given in Fig. 8 and the tools used for aircan on the Raspberry PI 3 during these steps are described below.

Step-1: In the first stage, the ODROID module installed on Raspberry PI 3 was taken in the monitor mode. The "airmon-ng start wlan0" command was used for this. Then the, "iwconfig" command is used to see that the module is in monitor mode. Fig. 9 shows the output of this process.

```

root@root:~# airmon-ng start wlan0

PHY      Interface  Driver      Chipset
phy3     wlan0     rt2800usb   Ralink Technology, Corp. RT5572

(mac80211 monitor mode vif enabled for [phy3]wlan0 on [phy3]wlan0mon)
(mac80211 station mode vif disabled for [phy3]wlan0)

root@root:~# iwconfig
wlan0mon IEEE 802.11 Mode:Monitor Frequency:2.457 GHz Tx-Power=20 dBm
Retry short long limit:2 RTS thr:off Fragment thr:off
Power Management:off

eth0     no wireless extensions.
lo       no wireless extensions.
    
```

Fig.9. Odroid module in monitor mode

Step-2: In the second stage, aircan was launched with the command "airodump-ng wlan0mon". For the first test case, the 00:11:22:33:44:00 BSSID mac address and the fake SSID called FU_TEST are published. In this scenario, only the SSID name was made similar to the current FU_TEST SSID name used on the network. There is no security setting in the SSID that the attacker issued when there was WPA2 encryption in the current FU_TEST broadcast. In this kind of attacks, users prefer the easy way because there is no authentication process

and connects to open SSID. The attacker then starts to sniff the traffic of the connected users. For this scenario, the attacker is set to have bssid value 00:11:22:33:44:00 over the FU_TEST for SSID broadcast in the “hostapd-mana.conf” file

```
root@root:~# airodump-ng wlan0mon

CH 9 ][ Elapsed: 2 mins ][ 2019-03-27 02:23

BSSID            PWR Beacons  #Data, #/s  CH  MB  ENC  CIPHER AUTH  ESSID
00:11:22:33:44:00 -35    23         0  0  6  11  OPN             FU_TEST
4C:FA:CA:4A:E6:C0 -69    19         2  0  1  720 WPA2 CCMP MGT  FU_TEST
```

Fig.10. Result of the first test case

According to the algorithm proposed in this type of attack, Raspberry PI 3 accepts this unsecured broadcast, fake SSID. Then the Raspberry PI connects backbone to ssh and assigns the BSSID mac address of this fake broadcast to an unauthorized VLAN previously set in the backbone. Users will not be able to access the Internet even if they connect to the SSID of this fake AP.

Step-3: In the third stage, the fake SSID with the 00:11:22:33:44:00 BSSID mac address and the same security settings as the original SSID of FU_TEST is issued. The airodump-ng was then started with the “airodump-ng wlan0mon” command as shown in Fig. 11. This kind of attacks are used to sniff traffic without making users feel. For this scenario, the attacker has made the following settings in the “hostapd-mana.conf” file in the “mana-toolkit” tool.

```
interface=wlan0
bssid=00:11:22:33:44:00
driver=nl80211
ssid=FU_TEST
channel=6
wpa=2
wpa_key_mgmt=WPA-EAP
wpa_pairwise=TKIP CCMP
wpa_passphrase=AsecurePassword
```

```
root@root:~# airodump-ng wlan0mon

CH 10 ][ Elapsed: 12 s ][ 2019-03-27 03:18

BSSID            PWR Beacons  #Data, #/s  CH  MB  ENC  CIPHER AUTH  ESSID
00:11:22:33:44:00 -46     3         0  0  7  11  WPA2 CCMP MGT  FU_TEST
4C:FA:CA:4A:E6:C0 -66     2         0  0  1  720 WPA2 CCMP MGT  FU_TEST
```

Fig.11. Fake AP broadcast with the same security settings as the original SSID

In order to prevent such attacks, the network administrator needs a table with bssid mac addresses of the original broadcast access points. Raspberry PI compares the bssid mac address, which it sees when it sees the SSID broadcast with

located in the “mana-toolkit” tool. FU_TEST broadcast is made from channel 6. The screen output for this scenario is shown in Fig. 10.

the same name and the same security settings, with the previously prepared list. If the bssid mac address of the fake broadcast access point is not in the pre-prepared list, then this SSID is published from the fake AP. Then the Raspberry PI is connected to the backbone switch with ssh. Then the fake access point bssid mac address is assigned to an unauthorized VLAN which previously created on the backbone switch. After this stage, users will not be able to access the Internet even if they connect to the SSID of this fake AP.

Step-4: In the last scenario, the attacker can do the same with the original AP as the bssid mac address of the access point where the attacker does the fake broadcast as well as the SSID and security settings. Fig. 12 presents the output of this scenario. Also for this scenario, the change made to the file “hostapd-mana.conf.” In the “mana-toolkit” tool is given below.

```
interface=wlan0
bssid=00:11:22:33:44:00
driver=nl80211
ssid=FU_TEST
channel=7
wpa=2
wpa_key_mgmt=WPA-EAP
wpa_pairwise=TKIP CCMP
wpa_passphrase=AsecurePassword
```

```
root@root:~# airodump-ng wlan0mon

CH 9 ][ Elapsed: 4 mins ][ 2019-03-27 03:25

BSSID            PWR Beacons  #Data, #/s  CH  MB  ENC  CIPHER AUTH  ESSID
4C:FA:CA:4A:E6:C0 -54    71         10  0  7  720 WPA2 CCMP MGT  FU_TEST
```

Fig.12. Fake AP attack with same bssid, SSID and security settings

In such attacks, access points on the campus WiFi network saw this attack and did not allow such an attack against it. The output from the WiFi Controller used in AP management in the campus WiFi network is shown in Fig. 13. Many of the WiFi controllers of this kind of attacks are black-listed bssidmac address fake broadcast. Therefore, in the algorithm given in Fig. 8, a method for this type of attack has not been proposed.

Attack Detection Information (Spoofing attack)

MAC Address	Channel	RSSI (dBm)	Monitor AP	Last Discovered At
4cfa-ca4a-e6c0	7	-79	ap-1	2019-03-27 10:25:56

Fig.13. WiFi controller attack log

The pseudo-code generated for the algorithm used in the proposed method is given below.

Start

Input: Sniff air for 802.11 beacons

Output: Predicted Fake AP

If there is more than one ssid with the same name

if SSIDs have the same security settings

Compare BSSIDs in the AP list and the beacon mac addresses you monitor with Raspberry PI

If the beacon mac address is in the AP list

Then it is Authorized AP

Else

Then it is Fake AP, send warning message and take Fake AP to the unauthorized VLAN

Else

Then it is Fake AP, send warning message and take Fake AP to the unauthorized VLAN

else

Then Sniff air again

End

III. CONCLUSION AND FUTURE WORKS

The use of wireless WiFi devices is increasing day by day with the widespread use of internet. As a result, the possibility of capturing the information flow via wireless signals is also increasing due to the existing security vulnerabilities and intriguing by the attackers. The ability of attackers to view, receive and modify user data using these vulnerabilities has become inevitable in internet areas where security measures are not taken sufficiently. In this study, a method is proposed against fake AP attacks which is one of the attack types in wireless WiFi networks. The scenarios of the proposed method and the types of attacks that can be made are determined and the success of the proposed method against these scenarios is emphasized. These attacks were attempted to be identified with the given algorithm and the pseudo-code snippet, and then the studies to prevent these attacks were presented. Kali Linux operating system has been installed on Raspberry PI 3 which is one of the most widely used SBC devices. It was also used with the Odroid module SBC to monitor wireless broadcasts. As a result of the study, it was seen that fake AP attacks were successfully detected and prevented by the proposed method.

In future studies, an application will be developed to find fake APs that have the same security settings as the same beaconmac address in the same SSID. This application will be run on raspberry PI or similar SBC device and when it sees more than the same SSID, these SSIDs will be connected to the engine and traceroute will determine the number of hop and the software will be prepared with an interface where the accesspoint is fake and reported.

ACKNOWLEDGMENT

This work was supported by the FUBAP (Firat University Scientific Research Projects Unit) under Grant No: TEKF.18.13.

REFERENCES

- [1] C. Xu, W. Jin, X. Wang, G. Zhao, and S. Yu, "MC-VAP: A multi-connection virtual access point for high performance software-defined wireless networks," *J. Netw. Comput. Appl.*, vol. 122, pp. 88–98, 2018.
- [2] D. Liu, B. Barber, and L. DiGrande, Cisco CCNA/CCENT exam 640-802, 640-822, 640-816 preparation kit. 2009.
- [3] V. Kumkar, A. Tiwari, P. Tiwari, A. Gupta, and S. Shrawne, "Vulnerabilities of Wireless Security protocols (WEP and WPA2)," *Int. J. Adv. Res. Comput. Eng. Technol.*, vol. 1, no. 2, pp. 2278–1323, 2012.
- [4] H. R. Hassan and Y. Challal, "Enhanced WEP: an efficient solution to WEP threats," 2005, pp. 594–599.
- [5] R. Heartfield et al., "A taxonomy of cyber-physical threats and impact in the smart home," *Computers and Security*. 2018.
- [6] S. Wong, "The evolution of wireless security in 802.11 networks: WEP, WPA and 802.11 standards," sans.org/rr/whitepapers/wireless/1109. php Retrieved, pp. 1–10, 2003.
- [7] S. Vibhuti, "IEEE 802.11 WEP Wired Equivalent Privacy Concepts and Vulnerability," *San Jose State Univ.*, no. Iv, 2008.
- [8] A. H. Lashkari, R. S. Hosseini, and F. Towhidi, "Wired equivalent privacy (WEP)," in *Proceedings - 2009 International Conference on Future Computer and Communication, ICFCC 2009*, 2009, pp. 492–495.
- [9] Y. Liu, Z. Jin, and Y. Wang, "Survey on security scheme and attacking methods of WPA/WPA2," 2010 6th Int. Conf. Wirel. Commun. Netw. Mob. Comput. WiCOM 2010, pp. 1–4, 2010.
- [10] A. H. Adnan et al., "A comparative study of WLAN security protocols: WPA, WPA2," in *Proceedings of 2015 3rd International Conference on Advances in Electrical Engineering, ICAEE 2015*, 2016, pp. 165–169.
- [11] J. Z. Liu Yong-lei, "Distributed method for cracking WPA/WPA2-PSK on multi-coreCPU and GPU architecture," no. November 2013, pp. 723–742, 2009.
- [12] S. Gold, "Cracking wireless networks," *Netw. Secur.*, vol. 2011, no. 11, pp. 14–18, 2011.
- [13] Y. Wang, Z. Jin, and X. Zhao, "Practical defense against WEP and WPA-PSK attack for WLAN," in 2010 6th International Conference on Wireless Communications, Networking and Mobile Computing, WiCOM 2010, 2010.
- [14] K. Bicakci and B. Tavli, "Denial-of-Service attacks and countermeasures in IEEE 802.11 wireless networks," *Computer Standards and Interfaces*, vol. 31, no. 5. pp. 931–941, 2009.
- [15] J. Bellardo and S. Savage, "802.11 Denial-of-Service Attacks: Real Vulnerabilities and Practical Solutions," in *USENIX security*, 2003, pp. 15–28.
- [16] X. Zha and M. Ma, "Security improvements of IEEE 802.11i 4-way handshake scheme," in 12th IEEE International Conference on Communication Systems 2010, ICCS 2010, 2010, pp. 667–671.
- [17] Z. Bai and Y. Bai, "4-Way handshake solutions to avoid denial of service attack in ultra wideband networks," in 3rd International Symposium on Intelligent Information Technology Application, IITA 2009, 2009, vol. 3, pp. 232–235.
- [18] S. H. Eum, Y. H. Kim, and H. K. Choi, "A Secure 4- Way Handshake in 802.11i Using Cookies.pdf," vol. 2, no. 1, 2008.
- [19] A. Alabdulatif, X. Ma, and L. Nolle, "Analysing and attacking the 4-way handshake of IEEE 802.11i standard," in 2013 8th International

Conference for Internet Technology and Secured Transactions, ICITST 2013, 2013, pp. 382–387.

[20] Internet, “4 Way Handshake.” .

[21] T. D. Nguyen, D. H. M. Nguyen, B. N. Tran, H. Vu, and N. Mittal, “A lightweight solution for defending against deauthentication/ disassociation attacks on 802.11 networks,” Proc. - Int. Conf. Comput. Commun. Networks, ICCCN, pp. 185–190, 2008.

[22] K. El-Khatib, “Impact of feature reduction on the efficiency of wireless intrusion detection systems,” IEEE Trans. Parallel Distrib. Syst., vol. 21, no. 8, pp. 1143–1149, 2010.

[23] K. F. Kao, W. C. Chen, J. C. Chang, and H. Te Chu, “An accurate fake access point detection method based on deviation of beacon time interval,” in Proceedings - 8th International Conference on Software Security and Reliability - Companion, SERE-C 2014, 2014, pp. 1–2.

[24] M. K. Chirumamilla and B. Ramamurthy, “Agent based intrusion detection and response system for wireless LANs,” 2004, pp. 492–496.

[25] S. Nikbakhsh, A. B. A. Manaf, M. Zamani, and M. Janbeglou, “A novel approach for rogue access point detection on the client-side,” in Proceedings - 26th IEEE International Conference on Advanced Information Networking and Applications Workshops, WAINA 2012, 2012, pp. 684–687.

BIOGRAPHIES



İlhan Firat Kilincer was born in Elazig, Turkey in 1986. He received B. S. degree in Electric Electronic engineering from Kocaeli University, Kocaeli, Turkey, in 2012. He is currently pursuing the Ph.D. degree in electrical engineering. He works as network security expert in the Computer

Center, Firat University, Elazig



Fatih Ertam was born in Elazig, Turkey in 1978. He received the B.S. and M.S. degrees in computer science from Firat University, in 2000 and 2005, respectively, and the Ph.D. degree in software engineering from Firat University, in 2016. He is currently an

assistant professor with the Digital Forensics Engineering, Firat University. His current research interests include network security, machine learning, deep learning, and network forensics.



Abdulkadir Sengur graduated from the department of Electronics and Computer Education at Firat University in 1999. He obtained his M.S. degree from the same department and the same university in 2003. His Ph.D. degree was from the department of Electronic Engineering at

Firat University in 2006. He is a professor at Firat University. His interest areas include pattern recognition, machine learning and image processing.

On Base Station Localization in Wireless Sensor Networks

H. E. KIZILOZ

Abstract— Wireless sensor networks (WSN) has been a prominent topic for the past decade. WSN consist of multiple sensor nodes, which collect and convey data to the base station(s). Sensor nodes are expected to run on batteries, and it makes energy the scarce resource for sensor nodes. The energy expenditure of a sensor node mainly depends on data transmission, which is exponentially affected by transmission distance. Consequently, if sensor nodes forward their data to the base station directly, distant sensor nodes will exhaust quickly. On contrary, minimization of transmission distance for each sensor node, i.e., each node transmits its data to the closest sensor node on its path to the base station, depletes the energy of sensor nodes that are closer to the base station fast. As a result, the flow balance in the network must be optimized. In this study, we investigate the effect of optimization of the base station location along with flow balance optimization. For this purpose, we compare five different localization methods on different topologies; three statically located linear programming approaches, a dynamically located nonlinear programming approach and a heuristic-based hybrid approach. Experimental results indicate that lifetime improvement of up to 42% is possible in selected scenarios.

Index Terms—lifetime optimization, linear programming, wireless sensor networks.

I. INTRODUCTION

WIRELESS SENSOR NETWORKS (WSN) has gained much attention recently. The emergence of cheap devices and the functionalities it provides make WSN an interesting field of topic. Even though military and/or emergency applications were the main causes in its appearance, it has become a common concept in academia, commercial market, and even among hobbyists [1]. Anyone can deploy their WSN at their homes, offices, etc. as of today [2]. Moreover, insights on the Internet of Things and urban computing raise interest in WSN [3].

A typical WSN consists of multiple sensor nodes and at least one sink called the base station. A sensor node is responsible for two actions: i) sensing, which is collecting data using its sensor. Any type of sensors may be used for data collection, e.g. humidity, air temperature, pressure, acoustic, magnetic, etc. ii) conveying its data to the base station. For this purpose,

sensor nodes may directly communicate with the base station, or they may choose to use other sensor nodes as a relay.

A possible drawback of WSN is that the number of sensors in a WSN topology could increase up to hundreds or thousands, and all sensor nodes are assumed to work on a battery, hence, they have limited and nonrenewable energy. Since energy is the scarce resource for a sensor node, the energy expenditure of the nodes must be optimized. Even though energy dissipation for sensing may be application or sensor type-specific; in general, a sensor node uses most of its energy for data transmission. As a result, minimization of the amount of energy used for data transmission helps to maximize the lifetime of the WSN.

Energy usage for data transmission is highly bound to the distance between communicating parties. Amount of energy required to transmit one-bit data increases exponentially as distance increases. However, selecting the closest sensor node as the relay causes the sensor nodes that are closer to the base station to exhaust all of its energy very quickly. Since we define the lifetime as the period until the first sensor node depletes its whole energy, the lifetime would be short in such a case. As a result, flows between sensor nodes must be optimized in such a way that it does not exhaust the nodes that are either close to or far from the base station.

Given the sensor node and base station locations, optimizing the flows between parties has been a resolved issue [4]. Since distance is the key factor for energy expenditure, there have been several studies for optimizing sensor node locations in literature [5, 6, 7]. The mentioned studies proposed different methods for sensor node deployment that increases lifetime and ensures coverage.

In this study, we look at the problem from a different angle. The main purpose of this study is to investigate the effects of using different base station localization methods over the lifetime in the WSN domain. For this purpose, we utilize the model proposed in [4] and compare five base station localization methods, in terms of their effect on the lifetime of the network and their execution times, with empirical results.

The rest of the paper is organized as follows. The literature review is given in Section 2. The model is described in Section 3, and experimental results are given in Section 4. We conclude the paper with concluding remarks in Section 5.

HAKAN EZGI KIZILOZ, is with Department of Computer Engineering University of Turkish Aeronautical Association, Ankara, Turkey, (e-mail: hakanezgi@etu.edu.tr).

 <https://orcid.org/0000-0002-2306-6008>

Manuscript received August 29, 2019; accepted January 03, 2020.

DOI: 10.17694/bajece.613154

Manuscript received August 12, 2018; accepted Nov 16, 2018.

DOI: [10.17694/bajece.419538](https://doi.org/10.17694/bajece.419538)

II. RELATED WORK

In this section, related work about base station location optimization in literature is given.

Oyman and Ersoy [8] proposed an iterative search for finding the optimal number and locations of base stations. After the sensor nodes are deployed, their model initiates with one base station. At each iteration, they apply the k-means clustering algorithm to locate the base stations at the center of each cluster and estimate a lifetime value accordingly. Iterations continue to increase the number of the base station until the estimated lifetime reaches a predetermined threshold value. Kim *et al.* [9] proposed a mixed-integer linear programming model for optimal multi-sink positioning and routing. In their model, sensor nodes are located into a topology first, and then some of them are selected as base stations during the optimization phase. Experiment results indicate improvement in lifetime and fairness when compared to multi-sink aware Minimum Depth Tree model.

In 2010, Güney *et al.* [10] proposed mixed-integer linear programming models for optimal base station location and routing in heterogeneous WSN for two separate objectives: minimizing energy expenditure, and minimizing financial cost. Also, they provided such techniques to find lower bounds. Later in 2012, the authors proposed a hybrid approach utilizing mixed-integer linear programming and Tabu Search to search for the optimal base station location [11]. In both models, sensor nodes and the base station(s) can only be located at predefined locations. Random deployment of the sensor nodes is also investigated in the latter study [11]. Different heuristic approaches also exist. Tripathi *et al.* [12] proposed a heuristic approach for finding the optimal base station location in two-tiered WSN. In their approach, they begin from the centroid of the nodes, assign link weights with respect to distances, and deploy the base station to the weighted average of the sensor nodes. Experimental results indicate that their proposed algorithm improves lifetime, yet the authors state that it does not find the theoretically optimal location. Fouad *et al.* [13] proposed a Particle Swarm Optimization (PSO) based heuristic method for finding the optimal base station location. According to experimental results, their method improves the performance of both topology construction and maintenance phases, and hence improves lifetime. Pan *et al.* [14] proposed a compact PSO method to find the optimal base station location. This method approximates the genetic algorithm and conventional PSO in terms of obtained results, but it requires less resources.

The work by Tung and Binh [15] is the most similar to this study. They use the same mathematical model, and they also compare five different base station localization methods. Their proposed method, integrated greedy method, improves the lifetime by 10%. Similarly, a recent study by Saha *et al.* [16] shows that their probabilistic localization method performs well.

The mobility of the base stations has attracted interest recently. Cayirpınar *et al.* [17] analyzed possible mobility patterns for multiple base stations, whereas Cicek *et al.* [18] provided mathematical formulations for flying base stations.

III. MODEL

In this section, we describe the employed energy model and base station localization methods separately, and in the specified order, in the following subsections.

A. Energy Model

In this study, we utilize the linear programming (LP) based energy model described in [4]. The model takes positions of the sensor nodes and the base station as input and optimizes the flow balance to maximize the lifetime of the network accordingly.

The energy model used in this study is as follows:

Maximize t

Subject to:

$$f_{ij} \geq 0, \forall (i, j) \in A \quad (1)$$

$$f_{ij} = 0, \forall (i, j) \in A, \text{ if } i \text{ is } B \text{ or } i = j \text{ or } d_{ij} \geq d_{max} \quad (2)$$

$$\sum_{j \in VUB} f_{ij} = \sum_{j \in VUB} f_{ji} + s_i t, \forall i \in V \quad (3)$$

$$E_{rx} \sum_{j \in V} f_{ji} + \sum_{j \in V} f_{ij} E_{tx,ij} + f_{iB} E_{tx,iB} \leq energy_i, \forall i \in V \quad (4)$$

In this model, the objective is to maximize lifetime, t , with respect to the constraints given with equations (1) through (4). We would like to mention that, B represents the base station, whereas V and A represent all sensor nodes and all edges (links) in the network, respectively. Moreover, data transmission (flow) from node i to node j is shown with f_{ij} . Among the constraints, Eq. (1) indicates that all flows must be non-negative. Flow in the network can either be positive, indicating that flow between sensor nodes exists; or zero, indicating that there is no flow between sensor nodes. In Eq. (2), we introduce additional constraints on flow. According to these additional constraints: i) base station cannot transmit data, ii) a sensor node cannot transmit data to itself, and iii) a sensor node cannot transmit data to another sensor node that is located further than d_{max} meters, which is 82.92 m. Flow balance at each sensor node is guaranteed with Eq. (3), that is, each node must transmit all the data they receive and produce. In this constraint, s_i represents the amount of data produced by each sensor node at a unit time interval (1 bps). Finally, Eq. (4) is a constraint for energy. The total amount of energy used by each sensor node cannot exceed its initial energy. In this model, sensor nodes are only responsible for collecting and conveying data. Complex tasks such as data aggregation or compression is not a concern. Also, the energy used for sensing and/or sleeping is considered as negligible. As a result, they only use energy for transmitting and receiving information. Energy used for receiving one bit of information is constant, $E_{rx} = \rho$, and equal to the electronics energy (50nJ). On the other hand, sensor nodes spend a variable amount of energy to transmit one bit of information, which is calculated with respect to the distance of the receiving node or

base station: $E_{ix,ij} = \rho + \varepsilon d_{ij}^\alpha$. In this formula, ε represents the amplifier energy ($100\mu\text{J}$), d_{ij} represents the Euclidean distance between sensor nodes i and j , and α represents path loss exponent. The $E_{ix,iB}$ value in Eq. (4) is a special case of $E_{ix,ij}$, where the receiver is not a sensor node but the base station. Finally, the path loss exponent, α , is considered low ($\alpha=2$) on a free space propagation topology without any obstacles, but it increases ($\alpha=4$) as propagation gets harder due to physical obstacles, such as rock, building, etc.

B. Base Station Localization Methods

In this study, we compare five different base station localization methods. Three of these methods employ static approaches, meaning that the base station is localized initially, and the network's lifetime is maximized with the given energy model. Another method dynamically searches for the best location for the base station, that maximizes the lifetime. The remaining approach is hybrid, and it utilizes a heuristic search algorithm, namely Particle Swarm Optimization (PSO). All the methods are described in detail as follows.

center: This is a static approach. The base station is simply located at the center of the area of interest, e.g. at the center of a circular topology. This approach is common in the literature.

mean: This is, again, a static approach. The base station is located at the mean values of x - and y -coordinates of the sensor nodes. Intuitively, this method seems to be a simple yet effective improvement to localization at the center.

random: Third and the final static approach. The base station is located at a random location inside the area of interest. A possible scenario for this model could be deploying both sensor nodes and the base station into a danger zone by throwing them from an airplane. In this scenario, the base station could be designed as a more powerful computing device with large enough energy, and a GSM connector.

dynamic: This method introduces two new variables to the model: x - and y -coordinates (position) of the base station. Rather than statically locating the base station, this method searches for the optimal location for the base station that maximizes the network's lifetime. The introduction of these two variables and their effects on the energy model (dynamically calculating the Euclidean distance between each sensor node and the base station) make it more complex, hence, the energy model becomes a nonlinear programming (NLP) model. Unlike an LP model, an NLP model may get stuck in a local optimum and cannot ensure global optima. Moreover, an NLP model may require more time to solve than an LP model. Nonetheless, it may still be valuable in this domain and it is worth testing it as a possible method.

hybrid: This hybrid approach utilizes PSO, a heuristic search algorithm. PSO algorithm consists of several particles searching for the best solution, i.e. x - and y -coordinates (position) of the base station that maximizes the fitness value, the lifetime. Each particle represents a possible solution. Initially, the particles are initialized randomly. Then, the fitness value for each particle is calculated. The fitness value is calculated as similar to the static approaches: the particle provides the position of the base station, and the LP energy model maximizes the lifetime. Once fitness calculation for each particle is complete, the particle having the best fitness value is

determined and all particles begin searching towards the best particle. Fitness evaluation, the best particle's determination, and updating each particle's new search location is called a generation in PSO. The algorithm continues until a termination criterion, i.e. completing a predefined number of generations, is met. Finally, the overall best particle through all generations is reported as the solution of the algorithm.

To eliminate any bias, all base station localization methods are tested on equal terms; that is, all sensor nodes are initially deployed randomly in a circular topology, then the base station is localized separately for each method on this network, and finally, the energy model maximized the network's lifetime. As a result, any deviation in the lifetime solely depends on the location of the base station, which is only affected by the localization method.

IV. EXPERIMENTAL RESULTS

We use GAMS [19] to solve our proposed model. We use the energy model described in the previous section, with each sensor node having an initial energy of $2W$. However, we would like to note that, initial energy and energy dissipation for transmitting and/or receiving data is hardware and application-specific; hence may vary in different hardware or applications. As a result, obtained lifetime results should not be evaluated as exact timespan in seconds, but should be considered as a ratio. Hence, we give normalized result values, i.e., all lifetime values are divided by the maximum lifetime value of the same context.

The hybrid method employs the PSO algorithm, which requires the determination of algorithm-specific and execution parameters. Algorithm specific parameters are obtained from [13], as follows: inertia weight w is 0.8 , cognitive coefficient c_1 and social coefficient c_2 are 2 , and the random numbers r_1 and r_2 are selected in the range $[0,1]$. The two execution parameters, population size and the number of generations, are selected as 5 and 12 , respectively.

A sample topology depicting base station locations in all five localization methods are given in Fig. 1.

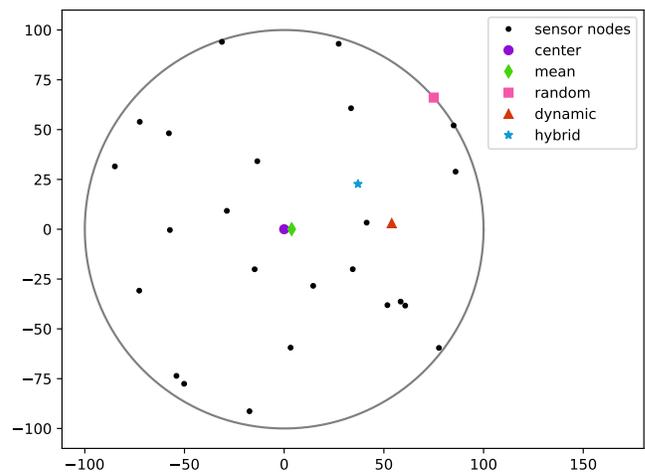


Fig.1. A sample topology with 25 sensor nodes, depicting base station locations for all localization methods.

In our experiments, we randomly deployed sensor nodes inside a circular topology having a radius of 100 m . We used varying numbers of sensors (25 , 50 , 100 , 250 and 500) to

simulate sparse and dense topologies. Also, we chose two different values for path loss exponent, $\alpha = 2$ and $\alpha = 4$, to simulate free space and multi-path propagation, respectively. Finally, to eliminate possible bias due to randomness, we share the average results of 1000 different sensor node deployments when the number of sensors is 25, 50 or 100. Topologies having 250 or 500 sensor nodes are averaged over a smaller number of deployments (100 and 30, respectively) due to their larger execution time (see Table 1).

Experimental results for the two path loss exponents, $\alpha = 2$ and $\alpha = 4$, are given in Fig. 2 and Fig. 3, respectively. In both figures, normalized lifetime values are plotted as the number of nodes increases. From the figures, it is clear that random deployment of the base station is a bad choice, for all cases. On free space propagation, where $\alpha = 2$, base station localization is more important for sparse topologies than it is for dense topologies. The dynamic method improves the lifetime for up to 7.1% in a sparse topology (where the number of sensor nodes is 25). On the other hand, as the network gets denser, the lifetime of all methods (except for the random method) approximates to the maximum. The hybrid method, contrary to its good lifetime improvement performance, is not a viable option due to its huge execution time. In topologies where multi-path fading is a concern, $\alpha = 4$, the huge execution time of the hybrid method may be negligible due to its fruitful solutions. In sparse topologies with multi-path fading, the hybrid method improves lifetime by 27.5%, 27.0% and 4.4% as compared to the center, mean and dynamic methods, respectively. Lifetime improvement becomes even more critical as the topology gets denser. The hybrid method improves lifetime by 42.6% as compared to the center and mean methods, and by 15.1% as compared to the dynamic method when the network consists of 500 sensor nodes.

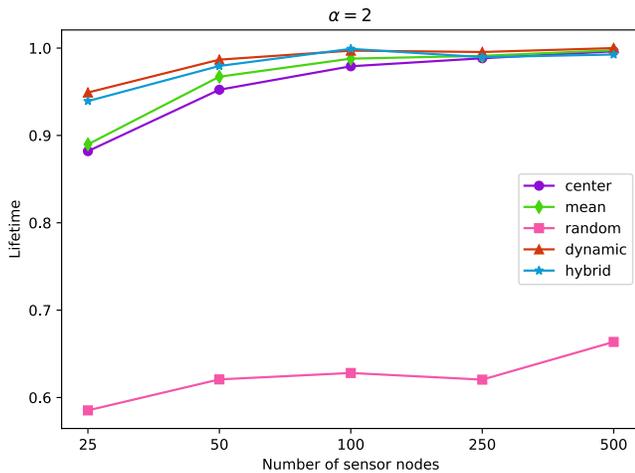


Fig.2. Normalized lifetime values with varying numbers of sensor nodes on free-space propagation topology ($\alpha=2$).

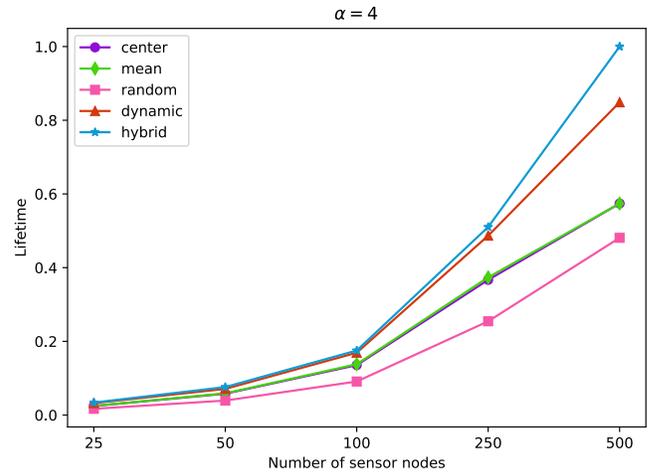


Fig.3. Normalized lifetime values with varying numbers of sensor nodes on multi-path propagation topology ($\alpha=4$).

We also provide average execution time for solving a problem in each model in Table 1. All static approaches require the same and minimum amount of the time, in all cases. The hybrid method requires roughly 60 times more amount of time as compared to the static methods, in all cases. This outcome is natural due to its implementation and the selected execution parameters (population size and the number of generations). In its comparison to the dynamic method, it requires more time in all cases, yet, the difference is negligible for dense topologies with multi-path fading. The dynamic method requires a similar amount of time to those of static methods for sparse and medium-sized topologies (up to 100 nodes). For dense topologies, however, its execution time increases massively.

Our findings on the lifetime and the execution times are summarized as follows.

- The base station should not be located randomly, if possible.
- On free space propagation, where $\alpha = 2$, the dynamic method can be executed in less than 5 seconds for sparse and medium-sized topologies, and it improves lifetime. For dense topologies, however, the base station can simply be located at the mean or the center.
- On multi-path propagation, where $\alpha = 4$, the hybrid method dominates all other methods in terms of improving lifetime. For sparse topologies, the dynamic method may still be a viable option considering its low execution time (under 5 seconds). For dense topologies, however, the hybrid method performs better and its usage is crucial.

TABLE I
AVERAGE EXECUTION TIMES OF SOLVING A PROBLEM FOR EACH TYPE OF TOPOLOGY AND LOCALIZATION METHOD (IN SECONDS).

# of nodes	$\alpha = 2$					$\alpha = 4$				
	25	50	100	250	500	25	50	100	250	500
center	0.55	0.57	0.75	3.92	40.14	0.52	0.57	0.66	3.38	37.59
mean	0.54	0.56	0.75	3.94	40.89	0.53	0.56	0.66	3.27	34.60
random	0.54	0.55	0.75	3.96	41.79	0.52	0.56	0.67	3.42	32.91
dynamic	0.59	0.81	3.55	91.53	992.60	0.56	0.87	4.12	185.30	2114.65
hybrid	30.55	32.13	46.45	235.68	2517.39	29.96	31.54	44.11	197.82	2120.15

V. CONCLUSION

Energy is the scarce resource of sensor nodes in wireless sensor networks (WSN), and it directly affects the lifetime of the network. The main factor in energy depletion is data transmission, which is highly dependent on transmission distance. Linear programming models have been employed so far to balance flows in a given WSN topology, to prevent premature energy depletion of a specific sensor node. These models work well, however, most of these studies assign a static location for the base station, e.g. the center. In this study, we analyze the effects of employing different base station localization methods on lifetime.

We evaluate five localization methods on different types of topologies: sparse or dense topologies on a free space channel or multi-path propagation. We compare the obtained results in terms of their effects on the lifetime and their execution time. Experimental results show that the dynamic method provides viable solutions in sparse and medium-sized topologies. It improves lifetime in a considerable amount of time (in under 5 seconds). In dense topologies, however, it becomes incompetent. Dense topologies having free space channel should employ the center or mean methods, whereas, the use of the hybrid method is crucial in dense topologies with multi-path propagation.

The main advantages and disadvantages of these methods are as follows. Among these methods, the random and center methods require no additional information and calculation of the base station location is imminent. The remaining methods require the knowledge of the sensor node locations, which may be hard or even impossible to acquire in a danger zone, e.g. natural disaster site. Moreover, the calculation time may be considered as a disadvantage, especially for the dynamic and hybrid methods. The most outstanding advantage is lifetime improvement, which is up to 42% in specific scenarios. Additionally, these methodologies are directly related, hence applicable to Internet of Things (IoT) studies, as WSN is a specialized form of it. Briefly, WSN may be considered as a more controlled and compact version of IoT, as IoT may consist of heterogeneous and geographically distributed devices.

Possible future work for this study could be making a similar analysis for networks having multiple base stations. Also, the effects of these methods on topologies combining different propagation channels could be investigated, e.g. $\alpha = 2$ for communication between some nodes, and $\alpha = 4$ for others.

REFERENCES

- [1] M. A. Gray and P. N. Scherer, "Web services framework for wireless sensor networks," in *SERVICE COMPUTATION 2014: The Sixth International Conferences on Advanced Service Computing, IARIA*, 2014, pp. 15–23.
- [2] S. Newberry and G. S. Gupta, "Wireless sensor based home automation as an educational springboard," in *Sensors Applications Symposium (SAS), 2015 IEEE*. IEEE, 2015, pp. 1–6.
- [3] B. Rashid and M. H. Rehmani, "Applications of wireless sensor networks for urban areas: A survey," *Journal of Network and Computer Applications*, vol. 60, pp. 192–219, 2016.
- [4] Z. Cheng, M. Perillo, and W. B. Heinzelman, "General network lifetime and cost models for evaluating sensor network deployment strategies,"

- IEEE Transactions on mobile computing, vol. 7, no. 4, pp. 484–497, 2008.
- [5] H. Subir, G. Amrita, S. Sanjib, D. Avishek, and D. Sipra, "A lifetime enhancing node deployment strategy in WSN," in *International Conference on Future Generation Information Technology*. Springer, 2009, pp. 295–307.
- [6] X. Liu, "Sensor deployment of wireless sensor networks based on ant colony optimization with three classes of ant transitions," *IEEE Communications Letters*, vol. 16, no. 10, pp. 1604–1607, 2012.
- [7] S. Singh, S. Chand, R. Kumar, and B. Kumar, "Optimal sensor deployment for WSNs in grid environment," *Electronics Letters*, vol. 49, no. 16, pp. 1040–1041, 2013.
- [8] E. I. Oyman and C. Ersoy, "Multiple sink network design problem in large scale wireless sensor networks," in *Communications, 2004 IEEE International Conference on*, vol. 6. IEEE, 2004, pp. 3663–3667.
- [9] H. Kim, Y. Seok, N. Choi, Y. Choi, and T. Kwon, "Optimal multi-sink positioning and energy-efficient routing in wireless sensor networks," in *International Conference on Information Networking*. Springer, 2005, pp. 264–274.
- [10] E. Güneş, N. Aras, İ. K. Altınel, and C. Ersoy, "Efficient integer programming formulations for optimum sink location and routing in heterogeneous wireless sensor networks," *Computer Networks*, vol. 54, no. 11, pp. 1805–1822, 2010.
- [11] E. Güneş, N. Aras, İ. K. Altınel, and C. Ersoy, "Efficient solution techniques for the integrated coverage, sink location and routing problem in wireless sensor networks," *Computers & Operations Research*, vol. 39, no. 7, pp. 1530–1539, 2012.
- [12] R. K. Tripathi, Y. N. Singh, and N. K. Verma, "Two-tiered wireless sensor networks - base station optimal positioning case study," *IET Wireless Sensor Systems*, vol. 2, no. 4, pp. 351–360, 2012.
- [13] M. M. Fouad, V. Snaesl, and A. E. Hassanien, "Energy-aware sink node localization algorithm for wireless sensor networks," *International Journal of Distributed Sensor Networks*, vol. 11, no. 7, 2015.
- [14] J. S. Pan, T. K. Dao, and T. S. Pan, "Compact particle swarm optimization for optimal location of base station in wireless sensor network," in *International Conference on Genetic and Evolutionary Computing*. Springer, 2016, pp. 54–62.
- [15] N. T. Tung and H. T. T. Binh, "Base station location-aware optimization model of the lifetime of wireless sensor networks," *Mobile Networks and Applications*, vol. 21, no. 1, pp. 10–17, 2016.
- [16] K. Saha, J. Aich, S. Chakraborty, and S. Bose, "Probabilistic Sink Placement Strategy in Wireless Sensor Network," in *Contemporary Advances in Innovative and Applicable Information Technology*. Springer, 2019, pp. 169–175.
- [17] O. Cayirpunar, B. Tavli, E. Kadioglu-Urtis, and S. Uludag, "Optimal mobility patterns of multiple base stations for wireless sensor network lifetime maximization," *IEEE Sensors Journal*, vol. 17, no. 21, pp. 7177–7188, 2017.
- [18] C. T. Cicek, H. Gultekin, B. Tavli, and H. Yanikomeroglu, "UAV Base Station Location Optimization for Next Generation Wireless Networks: Overview and Future Research Directions," in *2019 1st International Conference on Unmanned Vehicle Systems-Oman (UVS)*. IEEE, 2019, pp. 1–6.
- [19] G. D. Corporation, "General Algebraic Modeling System (GAMS) Release 24.2.1," Washington, DC, USA, 2013. [Online]. Available: <http://www.gams.com/>

BIOGRAPHIES



HAKAN EZGI KIZILOZ is currently an assistant professor at the University of Turkish Aeronautical Association. After receiving his BSc degree from the Mathematics department of TOBB University of Economics and Technology in 2008, he received his MSc and PhD degrees from the Computer Engineering department of the same university in 2010 and 2016, respectively. His research interests include usable security and optimization.

Desinging Analog Mixed Signal Circuits Using Graphene Nano Ribbon Field Effect Transistors

T. C. KARALAR

Abstract—Graphene is a 2D material formed by planar honeycomb placement of Carbon atoms. Besides its many superior physical properties it has superior electronic properties foremost of which is the high mobility it possesses. Due to this high mobility many Graphene based transistors have been designed. Graphene nano ribbons exhibit a band gap property, which is crucial for implementing transistors as switches. Moreover there exist models for these Graphene Nano Ribbon devices. In this work we first propose a gate array based manufacturing approach. Next assuming devices manufactured in this gate array methodology. We realize and simulate analog mixed signal blocks using Graphene Nano Ribbon transistors. The particular blocks that we used included telescopic amplifiers and StrongARM latches. Next we compared these blocks' performances against the same blocks implemented in 14nm high performance Silicon CMOS transistors. As a result we observed that the graphene transistors could attain comparable performances to circuits designed in 14nm CMOS. Specifically Graphene blocks can reach up to 80% of the bandwidth of Silicon devices. However Graphene devices have greater power consumption as a result of higher leakage current.

Index Terms—Graphene; Nano Ribbons; Nano Technology; Graphene Transistors; Graphene Circuit Design; Graphene Transistor modelling.

I. INTRODUCTION

GRAPHENE IS the 2D material comprised of a monolayer of carbon atoms arranged in a honeycomb lattice. Graphene and other 2D materials have superior electrical and mechanical properties than compared to other materials. For example, electrons can move relatively massless within this structure. As a result the electrons are delocalized and exhibit anomalous quantum Hall Effect. Moreover high electron mobility in Graphene, high Young's modulus and high thermal conductivity are all superior properties of this material that increase the possible uses of graphene. In short graphene is a multi purpose material that will have many industrial uses going forward [1-4].

High mobility is a very attractive electronic property of graphene. Due to this property same voltage can induce higher levels of current compared to other materials of similar dimensions. Therefore graphene transistors have many advantages for use as high-speed switching and amplifying devices.

TUFAN COŞKUN KARALAR, is with Department of Electronics and Communications Engineering University of Istanbul Technical University, Istanbul, Turkey, (e-mail: tufan.karalar@itu.edu.tr).

 <https://orcid.org/0000-0002-5424-0267>

Manuscript received August 05, 2019; accepted December 21, 2019.
DOI: [10.17694/bajece.601554](https://doi.org/10.17694/bajece.601554)

Despite these advantages transistors that can switch voltages and currents on and off has not been possible to realize. The main reason for this is the lack of a significant band gap between the valence and conduction bands in graphene devices. The electrons can move freely between valence and conduction bands. That is even though current can be created between the source and drain terminals of a transistor, this current cannot be reduced to significantly merely by changing the gate voltage. And the transistor on-current off-current ratio is not large enough. As a result graphene transistors using 2D graphene sheets as bodies cannot be used to design logic gates.

Single transistor amplifiers such as Low Noise Amplifiers or other single transistor circuits have been designed and published earlier. Also very simple analog circuits were simulated using GNRFET models.[5]

In this study using graphene transistor models developed for graphene nano ribbon based transistors are used to simulate practical analog mixed-signal circuits such as telescopic amplifiers and Strong ARM type latches. The novelty of our study comes from two aspects. First we are proposing a gate array based methodology to realize the multiple GNRFET devices. Second by using these devices and the GNRFET models available in literature we are simulating analog mixed signal blocks and estimating their performances. The existing GNRFET circuits are mainly digital designs [6, 7] however in our work we chose to study frequently used analog mixed signal circuit blocks such as telescopic amplifier and comparators. Additionally the same circuit blocks are designed using 14nm Silicon CMOS transistors. Finally two designs are compared in terms of a number of circuit parameters.

The rest of the paper is outlined as follows. Section II discusses the devices and intended gate array based manufacturing method. Section III describes the realized circuits. Section IV discusses the simulation results and comparisons between GNRFET and Silicon devices. Section V is the conclusion.

II. DEVICES

Circuits using 2D graphene transistors are generally used to design Graphene circuits. These circuits usually include single transistor LNA's, RF mixers and other amplifiers [8]. As mentioned earlier it's not surprising as Graphene devices are used to fast operation.

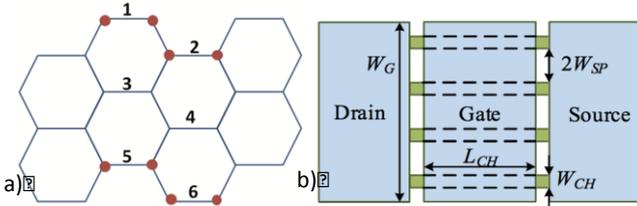


Figure 1 a) Graphene nano ribbon lattice diagram for a 6 dimer and armchair AGNR form structure [9] b) GNR transistor illustration. [10]

A. Physical Properties

When graphene is implemented in 1D it can have a larger band gap voltage. That is when graphene transistor bodies are realized on 10-100nm wide Graphene Nano Ribbons (GNR) a band gap voltage that is inversely proportional to the width [9, 10]. Despite the fact that below 100nm wide graphene nano ribbons exhibit band gap voltage when the graphene width goes over 100nm the band gap drops to converge eventually to the non-existent band gap observed on 2D graphene structures.

Due to the extreme dimensions of graphene nano ribbons we can assume that they would include a small number of dimer lines (5-20). Additionally we can assume that the ribbons will be in Armchair GNR configuration (AGNR) [10].

Illustrations of the physical and structural parameters of the Graphene Nano Ribbons as well as the GNR Field Effect transistors (GNRFET) are included in Figure 1. In this figure gate, source and drain terminals are assumed to be covering multiple of graphene nano ribbons, as transistor bodies. In this figure each terminal covers 4 nano ribbons

The cross section of a GNRFET is shown in Figure 2. Here a SiO₂ layer is added below the gate terminal. While the source and drain terminals are deposited right above the GNR sections.

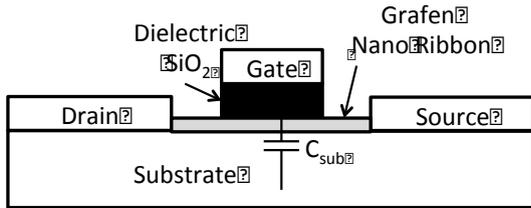


Figure 2 Transistor cross section

The source and drain terminals are expected to contact all the nano-ribbons. The distance between these two terminals constitute the channel length of the transistors. Nominally the channel lengths are assumed to be around 15nm. The total width of all the nano ribbons under the terminals defines the total width of the transistor. The transistor widths can be expected to assume discrete values, which is similar to FinFET devices.[11]. Using more transistors in parallel is the main mechanism to realize wider transistors. We assume that the nano-ribbons used in our case has 12-15dimer lines [12]. It can be noted that if the number of ribbons is low within the transistor the terminals need to be carefully aligned. However using a higher number of ribbons can alleviate any problems due to contact misalignment can be remedied.

In addition to the nano ribbon widths their separation is an additional physical parameter. This inter-ribbon spacing leads to a sub-optimal usage of area.

B. Fabrication Plan

Assuming a substrate with GNR's, and as seen in Figure 3 we can first shape GNRFET gate array [13] unit elements by etching graphene to create separations between GNRFET 's. Next regular arrays of electrodes are deposited to cover all GNR strips across all unit elements. For each transistor only 3 terminals are deposited. Two terminals are the source and drain electrodes while the third terminal is the gate electrode. The electrodes are anticipated as 5nm thick chromium layer and 100nm thick gold layer. The first two operations allow us to have a GNRFET transistor array and last but not least multi layer connections.

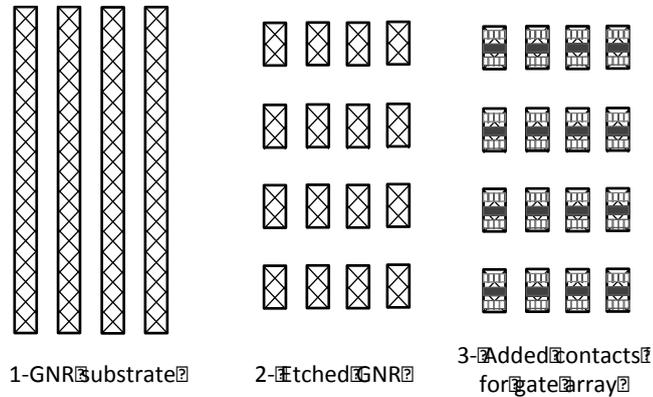


Figure 3 Preparation for GNRFET gate array

Once the devices and their electrodes are realized, the inter-device connections can be realized using SiO₂ insulation and gold metallization. In summary after fabricating a base device that includes structures similar to a gate array. The additional interconnect can be implemented to realize intended circuits. The final form of the intended circuit realization can be illustrated in Figure 4.

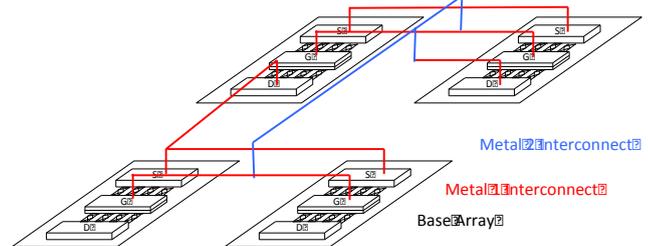


Figure 4 Gate array - like methodology.

The gate array implementation will obviously lead to inefficient use of area. For example to realize a transistor with W/L = 28n/14n we would roughly need 100n x 50n area per transistor. However till we can selectively grow GNR's the possible till then gate array methodology is a reasonable compromise.

C. Electronic Properties and Models

The transistor models that were utilized for this study are developed and published by UIUC [12]. They are compatible with SPICE format and can be used with commercial circuit simulators. The parameters used for these models are as follows: Channel length (L_{CH}), Nano ribbon width (W_{CH}), Gate terminal width (W_G), Spacing between ribbons (2W_{SP}), SiO₂ -oxide- thickness (t_{ox})

The main mechanism that creates the transistor current is the current induced within the channel [9]. The small dimensions of the transistors have significant impacts on their electronics parameters. First of all due to the small

dimensions of the devices their supply voltages are 0.8V. Therefore the any designs using these transistor models will need to observe low power and low voltage circuit design principles. Moreover again due to the small device sizes increased inter transistor capacitance and capacitive coupling will deserve extra attention.

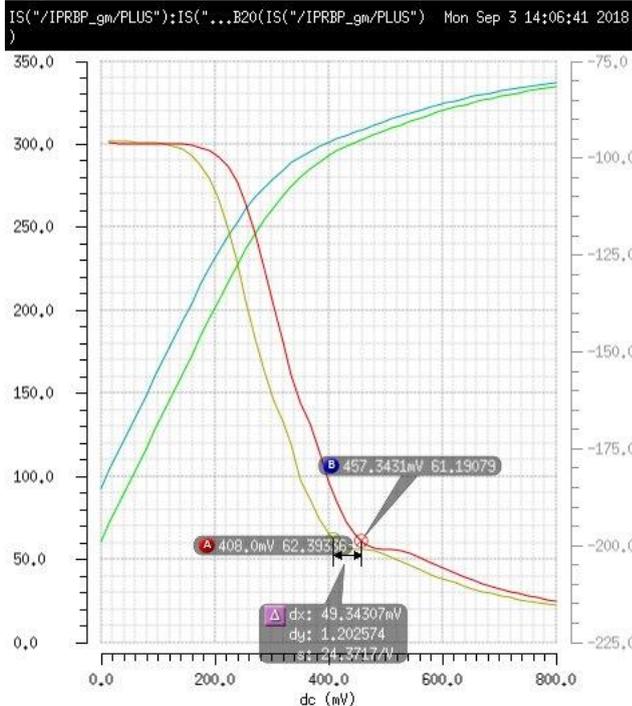


Figure 5. $\log(I_{DS})$ vs. V_{GS} Gate voltage for the GNRNFET (green) and GNRPFET (cyan) devices. The I-V curves' slopes for GNRNFET (red) and GNRPFET (yellow) are also plotted to estimate their threshold voltages of

Transistors use 0.8V as supply voltage. For transistors to operate successfully at such supply voltage levels the threshold voltages be accordingly. As a result of our simulations the NFET transistors using the Graphene Nano Ribbons (GNRNFET) have a threshold voltage of 457mV and PFET devices (GNRPFET) have thresholds of 408mV. The plot for the simulations supporting this conclusion is shown in Figure 5. Here Threshold voltage is defined as the inflection point of the logarithmic plot of the transistor current.

The internal voltage gain ($g_m * r_o$) of the transistors, which is an important criterion in analog circuit design, is simulated next. These results are shown in Figure 6 where the GNRPFET internal gain is at maximum 15 and the GNRNFET internal gain is at maximum 2100. When looked closer the NFET output resistance seems to have a value of 3 orders of magnitude larger than the PMOS.

III. REALIZED CIRCUITS

To analyze the performance of GNR based circuits mixed signal circuits will be designed using both GNRNFET transistor models and also using 14nm Silicon CMOS based transistor models. The performance results of these circuits will be compared to test the viability of GNRNFETs to be used in mixed signal circuits. The Silicon transistor are used assuming 14nm process node where the models are from Predictive Technology Modeling (PTM) variant which targets FinFET modeling [14]. Because the GNRNFET devices have 15nm channel length, the silicon transistors selected for comparison also had 15nm channel lengths.

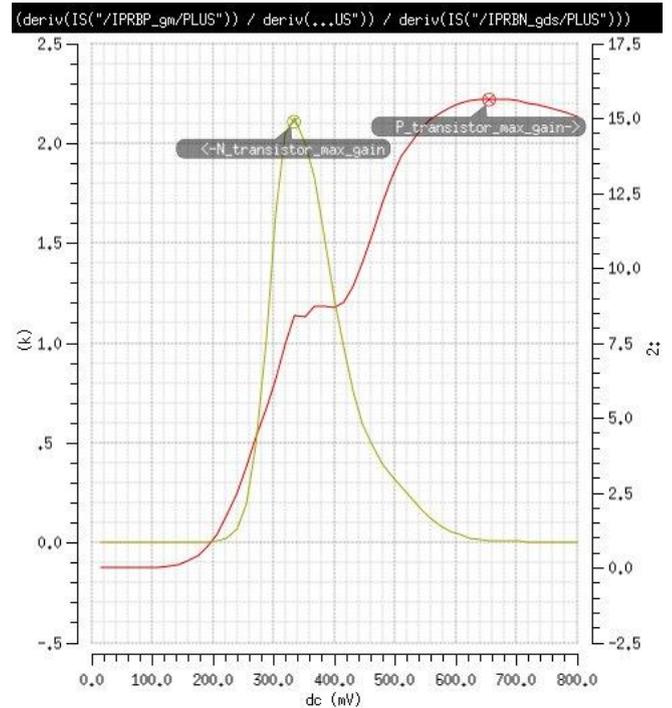


Figure 6 Internal voltage gains ($g_m r_o$) for GNRNFET (yellow) and GNRPFET (red) transistors.

To evaluate performance, we compared the speed, power consumption and area of similar circuits designed using GNRNFETs and Silicon transistors.

The first circuit used for comparison is a voltage buffer designed using a single stage telescopic amplifier [15] shown in Figure 7. This circuit was biased in 10uA and drives a 10fF load capacitance. The simulations were used to determine loop gain, bandwidth and voltage range. The second circuit is the important comparator circuit, which is common in many mixed signal circuits. In particular the popular StrongARM latches were used for the simulations [16]. The particular implementation is shown in Figure 8. The comparators are clocked at a 100MHz speed. Both these circuits are designed for comparison and they are compared in terms of their speed and power consumption.

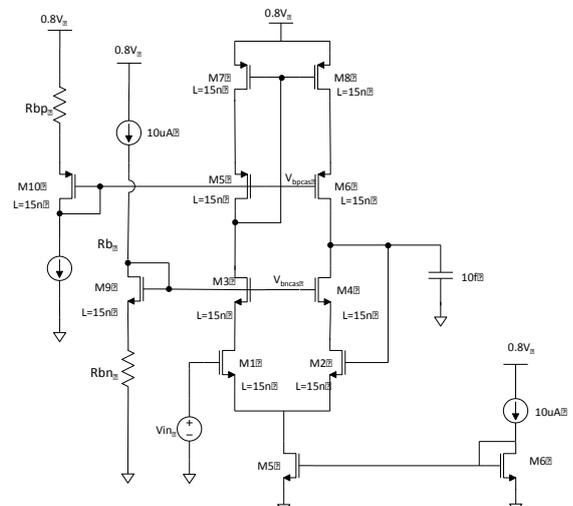


Figure 7 Voltage buffer designed using GNRNFET vs 14nm CMOS transistors.

Even though mismatch and noise performances are important for both circuits they were not included in our

study. The main reason for this is the models not including noise and mismatch information. This is because noise and mismatch parameters are measured as a result of statistical analysis of measured data. Therefore using the models in consideration it is not possible to complete noise and Monte Carlo Simulations.

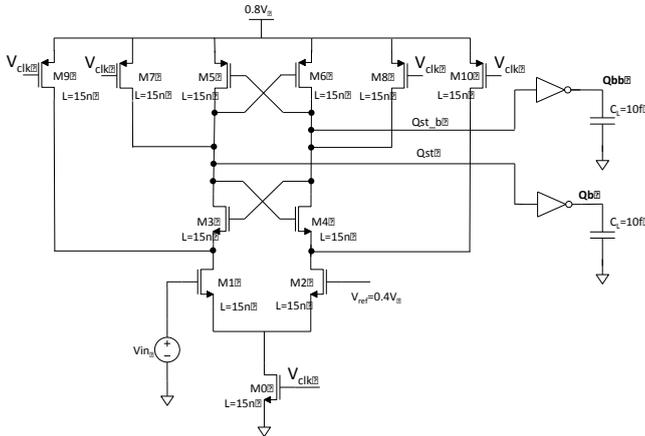


Figure 8 Comparator circuit implemented in GNRFET and 14nm CMOS devices

For area estimates an assumed 12nm S/D terminal widths are added onto the device channel lengths. Also 50%utilization is assumed so that the connections will increase the area consumption by an additional 50%.

IV. SIMULATIONS AND RESULTS

Tables 1 and 2 summarize the circuit performances obtained through simulations. Table 1 includes the performance results for the voltage buffer design. As can be seen from this table the same bias current consumption yields a 20% faster voltage buffer using Si CMOS transistors. One possible explanation for this advantage of Si FET devices over GNRFET devices can be the optimistic (high) mobility and (low) parasitic capacitance expectations in the devices assumed by the PTM models. Loop Gain DC value is 3dB larger for the Graphene implementation. This result is partly due to the unrealistically large DC gain of the GNRNFET devices. GNRPFET devices mainly dominate the DC gain of the graphene amplifier.

	Graphene	14nm Si
Bandwidth (GHz)	1.25	1.46
DC gain (dB)	29.89	27
Power Consumption(uA)	9.95	9.1
Area estimate (nm²)	50000	20000

Table 1 Comparison of voltage buffers designed using GNRFET and Si14nm FET devices

In terms of area consumption area estimate for both implementations the GNRFET's are implemented using gate arrays therefore have significant overhead. On the contrary the silicon devices can be built using custom layout and have reduced area overhead. As a result the area needed for Si circuits' are estimated to be less than half of the GNRFET implementations.

	Graphene	14nm Si
Output HL delay (ps)	60	61
Output LH delay (ps)	61	63
Power Consumption(uA)	6.6	2
Area estimate(nm²)	75000	30000

Table 2 Performance Comparison of comparators designed using GNRFET and Si14nm FET devices

As the result of the comparator simulations both implementations have a 60ps delay for the rising and falling edge transitions. The comparators are clocked at 100MHz speed. Meanwhile 14nm Si FET based implementations consume 60% lower average current. The current consumption is mainly during the regeneration phase when Vclk is high. The difference in power consumption is mainly due to the leakage current that the GNRNFET transistors exhibit when turned off. As mentioned before, we cannot complete comparator-offset analysis due to the lack of mismatch parameters.

V. CONCLUSION

In this study graphene nano ribbon transistors are used for mixed signal circuit design. Graphene based transistors have been used for design and simulations of logic gates and other digital circuits. On the contrary only single transistor amplifiers have been proposed in terms of analog mixed signal circuit design. To design the mixed signal circuits models for Graphene Nano ribbon based transistor are assumed. The attraction of these devices is the presence of a band gap present in 1D devices. With such a band gap nano ribbon based transistor can be used as electronic switches. As the mixed signal circuits to be used for comparison a voltage buffer and a comparator are implemented. These circuits are implemented using GNRFET transistors as well as 14nm Si FET devices. Comparing these implementations reveal that Si FET devices can realize 20% faster operation for voltage buffers while consuming speeds are comparable for StrongARM based comparators. Graphene power consumption is slightly larger. As a result graphene transistors can operate at comparable speed and power levels with their Si counterparts.

KAYNAKLAR

- [1] Y. Zhang, Y. Tan, H. Stormer, and P. Kim, "Fabrication of Graphene pnp junctions with contactless top gates," ed: Nature_London, 2005.
- [2] K. S. Novoselov, Z. Jiang, Y. Zhang, S. Morozov, H. L. Stormer, U. Zeitler, *et al.*, "Room-temperature quantum Hall effect in graphene," *Science*, vol. 315, pp. 1379-1379, 2007.
- [3] S. Pisana, P. M. Braganca, E. E. Marinero, and B. A. Gurney, "Graphene Magnetic Field Sensors," *IEEE Transactions on Magnetics*, vol. 46, pp. 1910-1913, 2010.
- [4] E. W. Hill, A. Vijayaraghavan, and K. Novoselov, "Graphene Sensors," *IEEE Sensors Journal*, vol. 11, pp. 3161-3170, 2011.
- [5] T. C. Karalar, "Grafen Nano Şeritlerden Oluşan Transistörleri Kullanarak Analog Devre Tasarımı," presented at the ELECO 2018 Elektrik Elektronik ve Biyomedikal Mühendisliği Konferansı, Bursa, Turkey, 2018.
- [6] H. V. R. Aradhya, H. R. Madan, M. S. Suraj, M. T. Mahadikar, R. Muniraj, M. Moiz, *et al.*, "Design, analysis and performance comparison of GNRFET based 8-bit ALUs," in *2016 International Conference on Research Advances in Integrated Navigation Systems (RAINS)*, 2016, pp. 1-6.
- [7] M. U. Mohammed and M. H. Chowdhury, "Design of Energy Efficient SRAM Cell Based on Double Gate Schottky-Barrier-Type GNRFET with Minimum Dimer Lines," in *2019 IEEE International Symposium on Circuits and Systems (ISCAS)*, 2019, pp. 1-4.

- [8] H. Madan, M. J. Hollander, J. A. Robinson, and S. Datta, "Analysis and benchmarking of graphene based RF low noise amplifiers," in *71st Device Research Conference*, 2013, pp. 41-42.
- [9] Y.-Y. Chen, A. Rogachev, A. Sangai, G. Iannaccone, G. Fiori, and D. Chen, "A SPICE-compatible model of graphene nano-ribbon field-effect transistors enabling circuit-level delay and power analysis under process variation," in *Proceedings of the Conference on Design, Automation and Test in Europe*, 2013, pp. 1789-1794.
- [10] M. Gholipour, Y.-Y. Chen, A. Sangai, and D. Chen, "Highly accurate SPICE-compatible modeling for single-and double-gate GNR-FETs with studies on technology scaling," in *Proceedings of the conference on Design, Automation & Test in Europe*, 2014, p. 120.
- [11] S. Sinha, G. Yeric, V. Chandra, B. Cline, and Y. Cao, "Exploring sub-20nm FinFET design with predictive technology models," in *Design Automation Conference (DAC), 2012 49th ACM/EDAC/IEEE*, 2012, pp. 283-288.
- [12] M. G. Ying-Yu Chen, Artem Rogachev, Amit Sangai, Deming Chen. (2014). *SPICE Model of Graphene Nanoribbon FETs*. Available: <https://nanohub.org/resources/17074>
- [13] M. Ashida, R. Hoshikawa, and J. Ishii, "CMOS Gate Arrays - Status, trends, design aids," in *ESSCIRC '82: Eighth European Solid-State Circuits Conference*, 1982, pp. 182-187.
- [14] Y. Cao. (2012, 09/04/2018). *Predictive Technology Models*. Available: ptm.asu.edu
- [15] B. Razavi, *Design of analog CMOS integrated circuits*. Boston, MA: McGraw-Hill, 2001.
- [16] B. Razavi, "The StrongARM Latch [A Circuit for All Seasons]," *IEEE Solid-State Circuits Magazine*, vol. 7, pp. 12-17, 2015.



BIOGRAPHIES

TUFAN COŞKUN KARALAR received his BS. Degree from University of Michigan, Ann Arbor in 2000, MS and PhD degrees from University of California, Berkeley in 2002 and 2006 respectively.

From 2006 to 2010 he was a staff design engineer working at Broadcom Corp. working on data converters, analog mixed signal circuits, filters, sensor interfaces. From 2010 to 2014 he was a Sr. Design engineer at Silicon Laboratories Inc. working on RF circuits, Analog mixed signal circuits, power electronics and digital circuit design. Since 2015 he has been an Asst. Prof at Istanbul Technical University Electronics and Communications Engineering Department. His research interests include, Sensor interface ASICs, High Speed Data converters and digital circuits, Power management IC's, Sensors in emerging technologies. He has 4 issued and 4 pending patents.

A New Bat Optimization Algorithm to Solve EPD Problem Solving with Transmission Loss

C. RAHEBI, M. AL-JUMAILI

Abstract—While, researchers work to make the systems operate economically and reduce operational cost, in this study, we work to reduce the fuel costs to make the power station running economically as much as possible by utilizing the Economic Power Dispatch (EPD) and the optimization algorithms. The economic power dispatch (EPD) is an integral part of the power system, The major roles and the purpose of its use are for achieving a reliable and efficient operation out of power system generation networks, and this operation should be obtained by minimizing the generator fuel cost. Getting optimal solutions to EPD problem requires efficient optimization algorithms. Novel Bat Algorithm (NBA) is one of the most recent methods and it has already proven its efficiency and reliability for solving the EPD problem. This paper proposes Novel Bat Algorithm (NBA) in order to solve the EPD problem based on the large scale power system. The NBA has proved its efficiency and it gave a perfect performance for small-scale systems compared with the original Bat algorithm (BA), because of considering the Doppler Effect and assumed that bat can move between various habitats. The study of the EPD for large-scale power systems is more important than study it for small ones, because all power systems in countries are considered large networks. This is why all modern studies focus on the study of EPD for large systems. To test the performance of NBA during small and large scale power system, we have applied it to many systems, including 3-thermal units, 6-thermal units, 31-Iraqi thermal units and IEEE 40-thermal units respectively, with transmission losses and generator limits, and we have made a comparison of the obtained results of NBA with other optimization methods. The Iraqi thermal units have been utilized as a new data.

Index Terms— Novel Bat Algorithm (NBA), Economic Power Dispatch (EPD).

I. INTRODUCTION

THE EPD is a main and an integral portion of power system and it is embedded under the term of economic operations of the power system. The purpose of EPD is to determine

CEVAT RAHEBI, is with Department of Electrical and Electronics Engineering Altinbas University, Istanbul, Turkey, (e-mail: cevat.rahebi@altinbas.edu.tr).

 <https://orcid.org/0000-0001-9875-4860>

MOHAMMED AL-JUMAILI, is with Department of Electrical and Electronics Engineering University of Turkish Aeronautical Association, Ankara, Turkey, (e-mail: eng_moh_86@yahoo.com).

 <https://orcid.org/0000-0002-3091-9638>

Manuscript received December 14, 2019; accepted January 30, 2020.
DOI: [10.17694/bajece.646870](https://doi.org/10.17694/bajece.646870)

optimal power outputs for the generating units in order to cover the load demand, according to the minimum fuel cost for each generating unit and satisfy different operational constraints over finite dispatch period [1]. That means, controlling the outputs for each generating unit is achieved by dealing with fuel cost equation for each thermal unit separately, and this status can be achieved by employing the optimization algorithms for the power system. As a result, the power system will start to operate economically and efficiently and the total fuel cost will be minimized and a lot of money can be saved. Because of that, we need to minimize the total fuel cost as much as possible, and it is required an efficient and reliable algorithm.

Many optimization methods were employed in order to solve the problem of EPD, and one of the most recent and efficient algorithms is the NBA. It was originally proposed in [2-4]. The proposed NBA has given a higher quality performance when it was applied for small scale power systems [3, 4]. The NBA performance can be achieved by making comparison with other well-known optimization methods that they proved their efficiency and reliability in solving the EPD problem, such as Genetic Algorithm (GA) [5-7], Particle Swarm Optimization (PSO) [8-11] and Quadratic Programming (QP) [12-14], and the performance of NBA will be studied by comparing the obtained simulation results of it with the other optimization algorithms which they mentioned above.

In this paper, two parts of dataset are used. The first part points to the small-scale power system and it consists of 3-thermal units and 6-thermal units. These networks have been taken from [3, 4]. The transmission loss is considered for these networks and it is calculated by the B-coefficient method. The second part of power network indicates to the large-scale power system and it also consists of 31-thermal units of Iraqi network and IEEE 40-unit system. Also the transmission loss is considered and it is calculated by the B - coefficient method. The study of the EPD for large-scale power systems is more important than the study of small-scale systems, because all power systems in countries are considered large networks. This is why all modern studies focus on the study of the EPD for large systems [15-17].

Many of optimization algorithms have been used to solve the problem of EPD problem and almost most of them have proven a good and high performance in reducing the fuel costs for running the electrical stations (thermal units) as we have shown previously, but the important goal that researchers are working on is to find the best algorithm which it will reduce fuel costs as much as possible in order to get an ideal

operation of generating units. In this study, we monitor the performance of NBA when it will apply to large scale systems, where the NBA has already proven that it is the best by applying it to small scale systems, and we also compare this performance with the other algorithms that we will show them later.

II. METHODOLOGY

A. Economic Power Dispatch

Economic Power Dispatch (EPD) shortly means the limitation of the values of the generators outputs for a specific power network, according to the lowest fuel cost for each unit, in order to cover the system load demand. EPD gives an efficient and economic operation of the power system by minimizing the fuel costs that are given by a specified combination of diverse generation units, while satisfying the units constraints, network equilibrium and load demand standard [18]. The EPD is a constrained and complicated optimization problem and it can be mathematically presented by minimizing the following equation:

$$F_T = \sum_{i=1}^n F_i(P_i) = \sum_{i=1}^n (a_i + b_i P_i + c_i P_i^2) \quad (1)$$

Where:

F_T : Total fuel cost of generation units (\$/HR), and F_i is the fuel cost function of the generator (i). The value (n) is the total generators number. (P_i) is the active power in (MW). The values of a_i , b_i and c_i are the fuel cost coefficients of generator (i).

The fuel cost in thermal units usually is taken as a function in quadratic form or other forms, and this function determines the total fuel cost for the unit according to the power output (P_i), and the total cost of fuel increases by increasing the power output of the generator as shown in Figure 1.

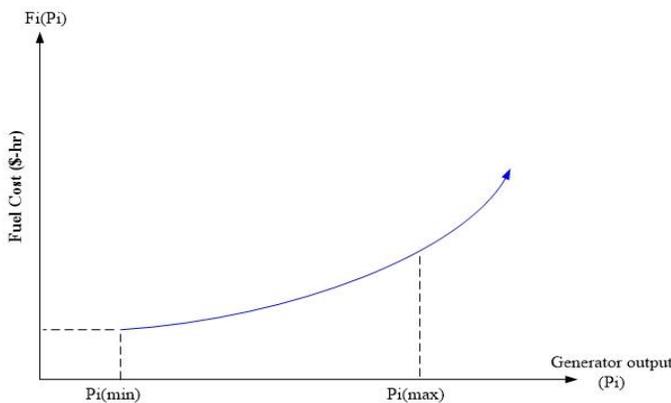


Fig. 1 Fuel cost characteristic of the thermal generating unit

The function of the fuel cost for a thermal generating unit can be expressed in quadratic form as follows:

$$F_i(P_i) = a_i + b_i P_i + c_i P_i^2 \quad (2)$$

B. Operational Constraints

There are some operational factors in the power systems,

and if we want to obtain optimal results in the study of EPD, these factors must be taken into consideration [19]. The operational constraints considered in our study can be presented as follows:

C. Power Balance Constraint

This constraint is classified as an equality constraint, and it can be achieved by equality the total power generation amount for all generating units in (MW) to the summation of total load demand of the system and transmission losses, and it can be expressed mathematically as follows:

$$P_D + P_L = \sum_{i=1}^n P_i \quad (3)$$

Where, P_D expresses the total load demand in (MW) and P_L expresses the transmission line losses in (MW) which they are calculated by using B-coefficients method and it can be presented mathematically as follows [20, 21]:

$$P_L = \sum_{j=1}^n \sum_{k=1}^n P_j B_{jk} P_k + \sum_{j=1}^n P_j B_{j0} + B_{00} \quad (4)$$

Where, $1 < j, k \leq n$ indicate to indexes of the generating thermal units, and B_{jk} , B_{j0} and B_{00} are known as the B-coefficients. This equality constraint is so important in the study of the EPD and it must be satisfied by the optimization algorithm.

D. Generation Capacity Constraint

The real power output for each generating unit must be specified within upper and lower limits and the obtained output of the generator must be during this limit, and it can be submitted as follows:

$$P_i^{min} \leq P_i \leq P_i^{max} \quad (5)$$

Where, P_i^{min} and P_i^{max} are the output powers of generators with maximum and minimum amount of generating unit (i).

E. Spinning Reserve Constraint

The aggregate of maximum power limits of whole generation units of a specified power system must be greater than the sum of total load demand and transmission losses. The mathematical equation of spinning reserve constraint generally can be expressed as follows:

$$\sum_{i=1}^n \sum_{L} P_{D_{imax}} \quad (6)$$

III. NOVEL BAT ALGORITHM

In the original BA, the effects of Doppler and foraging idea of bat were not considered and each essential bat is expressed by position and velocity and it searches preys during dimensional spaces with achieved trajectory[22, 23]. Actually, this case does not exist only. For the NBA, the Doppler effect is included and the essential bat can recompense adaptively for Doppler effect in phenomenon of

echoes [24].

The essential bat is viewed to possess foraging habitats diversely in the NBA. Bat searches for its food solely in one habitat in BA because of the mechanical conduct of virtual bat. In summary, the NBA is obligated for the following idealized basics:

1. The motion of bats can be around in various habitats.
2. Bats can recompense for the effect of Doppler in echo phenomenon.
3. Bats can acclimate and set their compensation averages and they depend on its target proximity.

A. Quantum Behavior

It is supposed that bats are going to conduct in such a behavior that while one of bats group found its prey in a specific habitat, immediately the rest bats would start to feed from the same prey. This supposition guides mathematically to the following formulation of the bat positions [24].

$$X_{i,j}^{t+1} = \begin{cases} g_j^t + \theta * |mean_j^t - x_{i,j}^t| * \ln\left(\frac{1}{u_{i,j}}\right), & \text{if } rand_j(0,1) < 0.5 \\ g_j^t - \theta * |mean_j^t - x_{i,j}^t| * \ln\left(\frac{1}{u_{i,j}}\right), & \text{otherwise} \end{cases} \quad (7)$$

B. Mechanical Behavior

It is supposed that the speed of the virtual bat will not overtake the sound speed which is estimated 340 m/s. The bat will compensate the Doppler Effect and this compensation will be expressed mathematically as CR that it varies among various bats. CR and the inertia weight (w) are in the range of 0 to 1. The value (ζ) represents the smallest constant to avoid the probability of division by 0. CR will be 0, if there is no compensation for the Doppler Effect by the bat, and it will be 1, if there is compensation. This description can be expressed mathematically as follows:

$$f_{i,j} = f_{minmax_{min}} \quad (8)$$

$$f_{i,j} = \frac{c+v_{i,j}^t}{c+v_{g,j}^t} * f_{i,j} * (1 + CR_i * \frac{g_j^t - X_{i,j}^t}{|g_j^t - X_{i,j}^t| + \zeta}) \quad (9)$$

$$V_{i,j}^{t+1} = w * V_{i,j}^t + (g_j^t - X_{i,j}^t) * f_{i,j} \quad (10)$$

$$X_{i,j}^{t+1} = X_{i,j}^t + V_{i,j}^t \quad (11)$$

C. Local Search

It is supposed logically that bats will raise the value of the pulse emission rate and reduce loudness when they approach prey. Whatever loudness value bats use, the loudness factor needs to be considered in the around environment. This description means the equations have been developed and expressed as follows:

If ($rand(0, 1) > r_i$)

$$X_{i,j}^{t+1} = g_j^t * (1 + tandn(0, \sigma^2)) \quad (12)$$

$$\sigma^2 = |A_i^t - A_{mean}^t| + \zeta \quad (13)$$

Where, $rand n(0, \sigma^2)$ is the Gaussian distribution with a mean value 0, (σ^2) is the standard perversion, and (A_{mean}^t) is the mean loudness.

In our study, we will use the NBA in [3] to check the performance of it based on large-large scale system. The parameters of the NBA are shown in Table 1, and the flowchart of it is shown in Figure 2.

Table I The parameters of NBA

Parameter	Value	Parameter	Value
Iterations (M)	1000	POP	30
rOMax, Min	0-1	A. MAX, MIN	1-2
f. D max, min	0-1.5	G	10
Prop.max, min	0.7-0.9	Theta max, min	0.5-1
C max, min	0.1-0.8	w. Max, min	0.5-.08

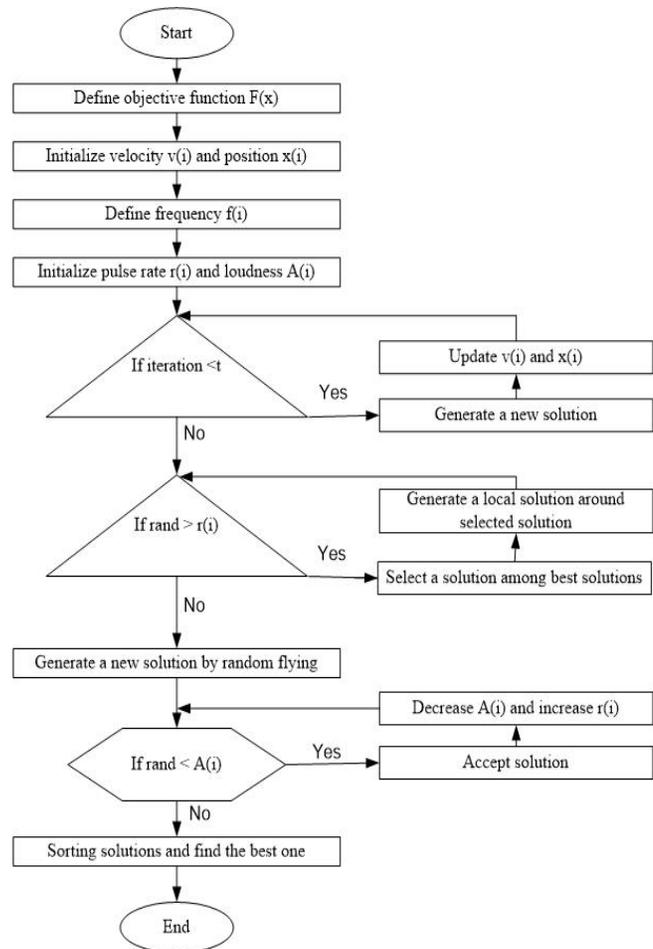


Fig. 2 Flowchart of NBA

IV. RESULTS AND DISCUSSION

The performance of the proposed NBA is tested with many power systems, including 3, 6, 31 and 40 generation units respectively. The intention of using several different networks is to monitor the performance of the algorithm during small and large scale system. The transmission line losses are done by the B-coefficient method. All the results are

calculated by using MATLAB R2015a.

A. Test of 3-unit system

In this test, three thermal units are utilized, and their cost coefficient and generator power limits are shown in [4]. The virtual system has been solved by the NBA and the achieved results of the NBA were compared with other optimization methods which they are GA, PSO and QP. The obtained results are shown in Table 2.

Table II Test results for 3-unit system with load demand 500 MW

Algorithm	P1 (MW)	P2 (MW)	P3 (MW)	Ploss	Cost (\$-hr)
NBA	107.3479	200.2140	219.5278	27.0897	26167.3565
GA	102.8798	189.7938	234.1524	26.8260	26177.6196
PSO	107.3453	200.2042	219.5406	27.0901	26167.3925
QP	107.4327	200.5192	219.1217	27.0736	28674.0000

From Table II, the assumed load demand was 500 MW and transmission losses were considered for this test and they have been calculated by the B - coefficient method. According to the obtained results, the NBA is the best in solving the EPD problem and it gave a higher quality performance in minimizing the fuel cost compared with the other algorithms.

B. Test of 6-unit system

This test has been done for six thermal units, and the cost coefficient and power limits of units are expressed in [3]. The achieved results have been solved by the NBA, and then there was comparison with other optimization methods which they are GA, PSO and QP.

From Table III, the suggested load demand was 700 MW for this test and the performance of NBA was effective in getting optimal results for generators output.

Table III Test results for 6-unit system with load demand 700 MW

Pi (MW)	NBA	GA	PSO	QP
P1	312.7083	286.6210	312.7083	312.7083
P2	72.5257	58.5750	72.5256	72.5252
P3	159.8879	171.0042	159.8878	159.8879
P4	50.0000	67.1123	50.0000	50.0000
P5	54.8817	62.3938	54.8816	54.8817
P6	50.0000	54.2969	50.0000	50.0000
Ploss (MW)	0.0036	0.0032	0.0033	0.0036
Cost (\$-hr.)	8299.4	8325.2	8299.418	8299.4

C. Test of 31-Iraqi unit system

The 31-unit system were utilized in this test as a large-scale power system and they are the Iraqi thermal unit. All details, including the cost coefficients and generator limits have been taken by the Iraqi ministry of electricity.

From Table 4 (A and B), the suggested load demand for this test was 4500 MW with considered transmission losses.

Table IV-A Test results for 31-unit system with load demand 4500 MW

Pi (MW)	NBA	GA	PSO	QP
P1	15	31.5749	15	15
P2	18.8356	55	15	15
P3	16.5991	55	15	15
P4	35.1213	160	35	35
P5	35.0037	160	35	35
P6	25.0161	138.0627	35	35
P7	35.0139	160	35	35
P8	300	165.1530	299.9999	300
P9	300	214.8092	299.9999	300
P10	300	300	299.9999	300
P11	300	300	299.9999	300
P12	220	34.0234	219.9999	220
P13	30.0288	220	219.9999	220
P14	220	220	219.9999	220
P15	220	220	219.9999	220
P16	220	220	219.9999	220

Table IV-B Test results for 31-unit system with load demand 4500 MW

Pi (MW)	NBA	GA	PSO	QP
P17	220	39.4235	219.9999	220
P18	20.5322	100	20	20
P19	100	100	20	20
P20	35.1448	48.7172	35	35
P21	35.0183	44.8555	35	35
P22	66.3707	94.2457	48.2087	48.2088
P23	76.2477	210	48.8043	48.8044
P24	72.9164	68.0822	48.0280	48.0281
P25	80.8614	46.2235	48.4702	48.4703
P26	610	311.2984	610	610
P27	610	122.2744	609.9999	610
P28	63.0292	171.2775	63	63
P29	63.1460	79.0213	63	63
P30	73.0908	218.7376	73	73
P31	73.0627	193.8328	73	73
Ploss	0.0389	1.6228	0.5115	0.5116
Cost\$-HR	60860	70012	60119.6	60131

From the obtained results, we found the performance of NBA started to decrease compared with the other algorithms that are shown below. The PSO was the best in getting an optimal solution.

D. Test of 40-unit system

In this test, IEEE 40-unit system (IEEE 30 bus system) has been utilized as a large-scale power system. The details of this system - including the cost coefficient and the power limits of generators are shown as a file on Math-Work site.

From Table 5 (A, B and C), the assumed load demand for this test was 6500 MW with considered transmission losses.

Table V-A Test results for IEEE 40-unit system with load demand 6500 MW

Pi (MW)	NBA	GA	PSO	QP
1	40	47.7988	40	40
2	60.3730	120	61.8501	61.8501
3	80.1652	81.4008	80	80
4	24.0472	42	24	24
5	28.3045	42	26	26
6	73.3975	140	68	68
7	111.4304	300	217.1984	217.1984
8	300	146.4729	258.7956	258.7956
9	135.0618	146.4857	260.0330	260.0330
10	130.0025	131.4314	130	130
11	94.3453	137.8951	94	94
12	94.0382	106.6941	94	94
13	195.3256	239.5136	195	195
14	196.0112	213.5743	195	195
15	195.2119	275.6720	195	195
16	195.0475	197.9386	195	195

Table V-B Test results for IEEE 40-unit system with load demand 6500 MW

Pi (MW)	NBA	GA	PSO	QP
17	195.1048	213.9538	195	195
18	220.0003	229.1696	257.2110	257.2110
19	388.8812	242.5194	260.0793	260.0793
20	550	469.9283	251.8288	251.8288
21	243.2754	284.6752	257.6483	257.6483
22	550	292.3266	494.6209	494.6209
23	254.1536	394.0592	495.2679	495.2679
24	550	274.9325	507.8255	507.8255
25	550	294.7605	514.2391	514.2391
26	550	550	446.9388	446.9388
27	254.0040	335.6258	447.6347	447.6347
28	10.2862	150	10	10
29	11.4445	15.1741	10	10
30	10.1169	14.3578	10	10

Table V-C Test results for IEEE 40-unit system with load demand 6500 MW

Pi (MW)	NBA	GA	PSO	QP
31	20.7115	70	20	20
32	20.0393	70	20	20
33	20.0050	22.6434	20	20
34	20.3573	23.4375	20	20
35	18.2462	20.5085	18	18
36	18.0036	18.1234	18	18
37	20.0263	31.0269	20	20

38	25.0251	25.3631	25	25
39	25.2612	31.4245	25	25
40	25.7147	60	25	25
Ploss	3.4205	2.9108	3.1719	3.1719
Cost (\$-HR.)	94865.73 6	110845.00 1	94020.383	94081.370

The achieved results showed that the NBA was not the best in getting an optimal solution. The PSO and QP were better than NBA at minimizing the fuel cost and getting optimal results.

Table 6 indicates to the summary of fuel costs obtained from the forth tests and these summarized results showed the performance of NBA compared with the obtained results of rest algorithms.

Table VI Summary of fuel cost

System	NBA	GA	PSO	QP
3-units	26167.3565	26177.6196	26167.3925	28674
6-units	8299.4	8325.2	8299.418	8299.4
31-units	60860	70012	60119.6	60131
40-units	94865.7363	110845.0015	94020.3830	94081.370

The NBA has given higher quality performance for test 1 and 2 (small-scale systems) and it recorded the best solution for EPD problem. But when NBA applied in test 3 and 4 (large-scale systems), it has recorded a low performance and accuracy compared with PSO, and the NBA will give a lower performance if the number of units increases. The PSO has given the best solution for EPD problem.

V. CONCLUSION

In this paper, we used the principle of Economic Power Dispatch (EPD) with optimization algorithms, which in turn, it reduces the operating costs of fuel that is used in thermal power plant, we proposed the NBA as an optimization method in order to solve the EPD problem. The aim of this study was to watch the performance of NBA with the large-scale power system. Four tests were used, including 3-units, 6-units, 31-units and IEEE 40-units respectively. The high quality performance of NBA has been already proved in small-scale power system such as 3 and 6 units. When NBA has been applied for large-scale power system, including 31 and 40 units, the performance of it started to decrease gradually in minimizing the fuel cost compared to other well-known optimization methods which they have been used and they were including Genetic Algorithm (GA), Particle Swarm Optimization (PSO) and Quadratic Programming (QP). According to the achieved results, we concluded that NBA performance with small-scale power systems is better than large ones. In addition, most tests proved that the GA algorithm was the best in minimizing the transmission line loss.

REFERENCES

- [1] F. Abu-Mouti and M. El-Hawary, "Novel constrained search-tactic for optimal dynamic economic dispatch using modern meta-heuristic optimization algorithms," in *Electrical Power and Energy Conference (EPEC)*, 2011 IEEE, 2011, pp. 170-175.
- [2] P. Sujatha and K. Anjaneyulu, "A novel approach for security constrained unit commitment using bat computation," in *Advance Computing Conference (IACC)*, 2015 IEEE International, 2015, pp. 623-628.
- [3] S. Gautham and J. Rajamohan, "Economic Load Dispatch using Novel Bat Algorithm," in *Power Electronics, Intelligent Control and Energy Systems (ICPEICES)*, IEEE International Conference on, 2016, pp. 1-4.
- [4] Y. A. Gherbi, H. Bouzeboudja, F. Lakdja, F. Z. Gherbi, and D. Ould-Abdeslam, "New approach for solving economic load dispatch problem," in *Electrical Sciences and Technologies in Maghreb (CISTEM)*, 2014 International Conference on, 2014, pp. 1-5.
- [5] F. K. Khosa, M. F. Zia, and A. A. Bhatti, "Genetic algorithm based optimization of economic load dispatch constrained by stochastic wind power," in *Open Source Systems & Technologies (ICOSST)*, 2015 International Conference on, 2015, pp. 36-40.
- [6] B. M. Ramadan, T. Logenthiran, R. Naayagi, and C. Su, "Hybridization of genetic algorithm and priority list to solve economic dispatch problems," in *Region 10 Conference (TENCON)*, 2016 IEEE, 2016, pp. 1467-1470.
- [7] R. S. Barros, O. A. C. Cortes, R. F. Lopes, and J. C. Da Silva, "A Hybrid Algorithm for Solving the Economic Dispatch Problem," in *Computational Intelligence and 11th Brazilian Congress on Computational Intelligence (BRICS-CCI & CBIC)*, 2013 BRICS Congress on, 2013, pp. 617-621.
- [8] S. H. Elyas, P. Mandal, A. U. Haque, A. Giani, and T.-L. B. Tseng, "A new hybrid optimization algorithm for solving economic load dispatch problem with valve-point effect," in *North American Power Symposium (NAPS)*, 2014, 2014, pp. 1-6.
- [9] D. Santra, A. Mukherjee, K. Sarker, and D. Chatterjee, "Hybrid PSO-ACO algorithm to solve economic load dispatch problem with transmission loss for small scale power system," in *Intelligent Control Power and Instrumentation (ICICPI)*, International Conference on, 2016, pp. 226-230.
- [10] Z. N. Zakaria, A. Azmi, M. S. Laili, S. A. S. Jamalil, and M. H. Sulaiman, "An Extension of Particle Swarm Optimization (E-PSO) Algorithm for Solving Economic Dispatch Problem," in *Artificial Intelligence, Modelling and Simulation (AIMS)*, 2013 1st International Conference on, 2013, pp. 157-161.
- [11] J. Zhu, *Optimization of power system operation vol. 47*: John Wiley & Sons, 2015.
- [12] F. Benhamida, I. Ziane, S. Souag, A. Graa, and B. Dehiba, "Solving dynamic economic load dispatch with ramp rate limit using quadratic programming," in *North American Power Symposium (NAPS)*, 2013, 2013, pp. 1-5.
- [13] R. Danaraj and F. Gajendran, "Quadratic programming solution to emission and economic dispatch problems," *Journal-Institution of Engineers India part el electrical engineering division*, vol. 86, p. 129, 2005.
- [14] D. Bisen, H. M. Dubey, M. Pandit, and B. Panigrahi, "Solution of large scale economic load dispatch problem using quadratic programming and GAMS: a comparative analysis," *Journal of Information and Computing science*, vol. 7, pp. 200-211, 2012.
- [15] F. Guo, C. Wen, J. Mao, J. Chen, and Y.-D. Song, "Hierarchical Decentralized Optimization Architecture for Economic Dispatch: A New Approach for Large-scale Power System," *IEEE Transactions on Industrial Informatics*, 2017.
- [16] F. Guo, C. Wen, and L. Xing, "A distributed algorithm for economic dispatch in a large-scale power system," in *Control, Automation, Robotics and Vision (ICARCV)*, 2016 14th International Conference on, 2016, pp. 1-6.
- [17] W. Zhao, M. Liu, J. Zhu, and L. Li, "Fully decentralised multi-area dynamic economic dispatch for large-scale power systems via cutting plane consensus," *IET Generation, Transmission & Distribution*, vol. 10, pp. 2486-2495, 2016.
- [18] D. P. Kothari and I. Nagrath, *Modern power system analysis*: Tata McGraw-Hill Education, 2003.
- [19] A. J. Wood and B. F. Wollenberg, *Power generation, operation, and control*: John Wiley & Sons, 2012.
- [20] J. Sun, V. Palade, X.-J. Wu, W. Fang, and Z. Wang, "Solving the power economic dispatch problem with generator constraints by random drift particle swarm optimization," *IEEE Transactions on Industrial Informatics*, vol. 10, pp. 222-232, 2014.
- [21] J. Li, Q. Pan, P. Duan, H. Sang, and K. Gao, "Solving multi-area environmental U+ 002F economic dispatch by Pareto-based chemical-reaction optimization algorithm," *IEEE/CAA Journal of Automatica Sinica*, 2017.
- [22] S. Biswal, A. Barisal, A. Behera, and T. Prakash, "Optimal power dispatch using BAT algorithm," in *Energy Efficient Technologies for Sustainability (ICEETS)*, 2013 International Conference on, 2013, pp. 1018-1023.
- [23] A. M. Taha and A. Y. Tang, "Bat algorithm for rough set attribute reduction," *Journal of Theoretical and Applied Information Technology*, vol. 51, pp. 1-8, 2013.
- [24] X.-B. Meng, X. Gao, Y. Liu, and H. Zhang, "A novel bat algorithm with habitat selection and Doppler effect in echoes for optimization," *Expert Systems with Applications*, vol. 42, pp. 6350-6364, 2015.

BIOGRAPHIES



CEVAT RAHEBI was born in Urmia, Iran, in January 1982. He received the B.S. degree in Communication Engineering in 2005 from the Faculty of Engineering Science at the Azad University of Urmia, Iran. He received the M.S. degrees in Communication System Engineering in 2009 from the Sadjad University of Technology, Mashhad, Iran. He received the Ph.D. at the Department of Electrical and Electronics Engineering, Gazi University, Ankara, Turkey. His current research interests lie in the field of optimization algorithms, communication engineering and wireless sensor networks.



MOHAMMED ALJUMAILI was born in Iraq, in December 1986. He received the M.S. degrees in Electrical & Electronics Engineering Department in 2017 from the University of Turkish Aeronautical Association, Ankara, Turkey. He is working in the Iraqi Ministry of Electricity. His current research interests lie in the field of optimization algorithms, electrical engineering and renewable energies.

A Diagnosis Of Stator Winding Fault Based on Empirical Mode Decomposition in PMSMs

R. SELÇUK, Z. DOGAN

Abstract— Stator winding faults may cause severe damages in Permanent Magnet Synchronous Motors (PMSM) if not detected early on. The earliest fault detection in motors should be made during transient states throughout the initial starting period. A new approach based on Empirical Mode Decomposition (EMD) and statistical analysis was presented for detecting stator winding fault by way of transient state phase current of PMSM in this study. Models based on finite elements method were developed for the PMSM representing the healthy and faulty states in order to implement the suggested fault detection method. Afterwards, transient state stator phase winding currents were measured for healthy and faulty states under nominal load in accordance with motor models. These non-linear current signals monitored were separated into its Intrinsic Mode Functions (IMF) via the EMD method. Pearson Correlation Coefficient was used for determining the IMF that most resembles the characteristics of the main signal. Statistical parameter-based feature extractions were carried out for the IMF signals determined for the healthy and faulty states. Fault and fault level detection were carried out successfully by comparing the obtained feature vectors. The acquired results have put forth that the suggested method can be used securely for fault detection in electrical machines especially for early fault detection.

Index Terms— Permanent Magnet Synchronous Motors, Stator winding faults, EMD.

I. INTRODUCTION

PERMANENT MAGNET synchronous motors (PMSMs) are frequently used in various industrial and military applications such as wind energy, electrical vehicles, railroad transportation, military planes etc. due to their various advantages such as high efficiency, high torque, power and flux density, wide speed ranges, long service life, simple structure and precise torque control [1–6]. As is the case for all electrical machines, various faults may develop in PMSMs

due to environmental factors, production and installation faults, long term use under adverse conditions, stator open circuit transformations, thermal, electrical or mechanical strains [2,7]. A fault that develops in motors may result in system shutdown. These shutdowns may lead to loss of production time and raw material. The faults may increase maintenance costs and result in power losses. A minor fault may advance further causing more significant faults [8,9]. Thus, early fault detection is of significant importance for reducing fault ratio and preventing possible problems.

PMSM faults are classified in three groups as mechanical faults, magnetic faults and electrical faults [10]. Statistics indicate that more than 47 % of the electrical motor faults are due to electrical faults [11]. Current studies indicate stator winding faults among the most frequently observed faults in electrical machines including PMSMs. Stator fault is due to the gradual decrease in the insulation between the stator winding threads [10, 12–14]. The fault detection is the process of detecting and identifying the type and scope of the fault. Typically, a detection process enables the perception or estimation of one or more machinery output such as current [15], [16], voltage [2], [17], vibration [7], [10], and torque [18] by way of sensors. The current signals carrying the fault information are frequently preferred in fault detection processes due to their low cost and ease of access.

Various signal processing methods are used for fault detection in electrical motors. Typical signal processing methods are time domain analysis, frequency domain analysis, time-frequency domain analysis, Park's vector approach [19]. Fast Fourier Transformation (FFT) which is an analysis method at the frequency domain is among the most frequently used signal processing analysis carried out using the harmonic components of the signal as fault feature (FFT) [20]. Since time data is lost during analyses via FFT signal processing method, it is only used for processing fixed signals and sometimes it is difficult to distinguish similar harmonics with this analysis. It is necessary to pass over to the time-frequency domain analysis for overcoming this shortcoming [10,20,21]. Various analysis methods such as Short-time Fourier transformation [22], wavelet transformation [15], Hilbert Huang transformation (HHT) [23], empirical mode decomposition (EMD) have been suggested for time-frequency domain analyses. EMD is an effective method among these other methods for linear and unstable signals [2, 24].

There are many studies on winding faults in literature J. Rosero et al. carried out a study analyzing short circuit faults in PMSMs. Intrinsic mode functions (IMFs) were calculated in

RUMEYSA SELÇUK, is with Department of Computer Engineering University of Beykoz University, Istanbul, Turkey, (e-mail: rumeysaselcuk@beykoz.edu.tr).

 <https://orcid.org/0000-0003-1085-095X>

ZAFER DOGAN, is with Department of Electrical Electronics Engineering University of Tokat Gaziosmanpasa, Tokat, Turkey, (e-mail: zafer.dogan@gop.edu.tr).

 <https://orcid.org/0000-0002-7953-0578>

Manuscript received November 25, 2019; accepted January 16, 2020.
DOI: [10.17694/bajece.650484](https://doi.org/10.17694/bajece.650484)

this study via EMD method using the stator current. Afterwards, HHT was applied for calculating the instantaneous frequency values of IMFs. In addition, IMF1 and IMF2 were analyzed using Wigner Ville Distribution (WVD). This study puts forth the capability for determining unstable fault current types for different speed intervals [24]. In their studies, Mejia-Barron et al. presented the application of empirical mode decomposition-based methods such as Ensemble EMD (EEMD) and complete EMD (CEEMD) for inrush current analysis. IMFs were used for focusing on acquiring data on the inter-turn faults in the transformer [25].

Steady state motor currents have been used in majority of the studies in literature in this field for determining PMSM winding faults [12,15]. The faster faults are detected in electrical machines, the lower the possible economic losses are. Early stage fault detection in motors is made by way of transient signals monitored during the initial starting time. Because transient durations are on the order of milliseconds. Motor transient current is the best data for this purpose [26–28].

The present study puts forth a detection method for stator winding fault in PMSM based on the EMD analysis of Transient Motor Current Signals (TMCSs). Initially, simulation models for the healthy and faulty states of the PMSM were developed in Ansoft Maxwell environment via Finite Elements Method (FEM) and TMCS values were measured under full load conditions. Afterwards, the TMCSs were analyzed via EMD and the IMFs of these currents were calculated. The Pearson Correlation Coefficient (PCC) was then used for determining the IMF that is closest to the main signal. Finally, the statistical characteristics of this determined IMF signal was used for early stage fault detection.

The remainder of this paper is organized as follows. Section 2 presents the FEM simulation model of PMSM for healthy and faulty situations. Section 3 describes the proposed signal analysis and feature extraction. The analysis results obtained in section 4 are given and interpreted in detail. Section 5 is the conclusions section.

II. PERMANENT MAGNET SYNCHRONOUS MOTOR

Finite elements method (FEM) is a method that enables the examination of the healthy and faulty states of an electrical motor in a computer environment by entering required parameters for that particular motor. It enables performance monitoring without harming the motor since it requires no experiments in a laboratory environment [17]. Maxwell equations are used to solve electromagnetic problems [25]. Thus, the FEM software uses Maxwell equations to reach the solution. Maxwell equalities are provided below [29].

Eq. (1) shows the relationship between electrical field, E and flux density, J .

$$J = \sigma E \quad (1)$$

The induced field expression is indicated by Eq. (2).

$$\nabla \times E = -\nabla \times A \quad (2)$$

Electrical field is expressed by Eq. (3) in the two-dimensional problem.

$$E = -(A \times \nabla V) \quad (3)$$

Eq. (4) is obtained by placing this electrical field equation in Eq. (1).

$$J = -\sigma A - \sigma \nabla V \quad (4)$$

Flux density expression is indicated by (5) according to magnetic vector potential and voltage value gradient.

$$\nabla \times \left(\frac{1}{\mu(B)} \right) \nabla \times A = -\sigma A - \sigma \nabla V + J_{source} \quad (5)$$

μ , denotes magnetic permeability of the environment, ∇V , represents the gradient for the voltage in the area outside the conductive material in two-dimensional magnetic field problems. As such, the amplitude and phase angle values of the value A are calculated.

In the present study, FEM analysis was used to generate a Maxwell 2D model of PMSM since it can put forth a model that is closest to the actual system. Table 1 presents the motor parameters of the designed model and the full model is shown in Fig. 1.

TABLE I
PMSM PARAMETERS DESIGNED WITH FEM

PMSM Parameters	
Frequency: 50 Hz	Stator External Diameter: 340 mm
Rated Voltage: 138 V	Rotor External Diameter: 224 mm
Output Power: 20 kW	Stator Size: 100 mm
Rated Speed: 1000 rpm	Stator Number of Slots: 36
Number of Poles: 6	Connection: Star

As can be seen from the full model of PMSM, the stator winding type for the motor is whole-coiled. It has a rotor surface mounted magnet with a magnet type of XG196/96.

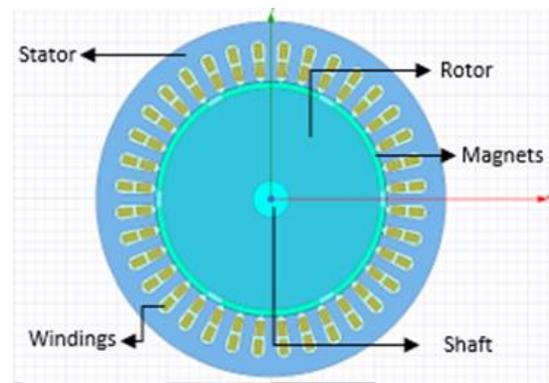


Fig.1. PMSM model with FEM

The 1/6 sampled section of the full model was analyzed in order to keep the analysis time shorter. Fig. 2 presents the PMSM analysis results.

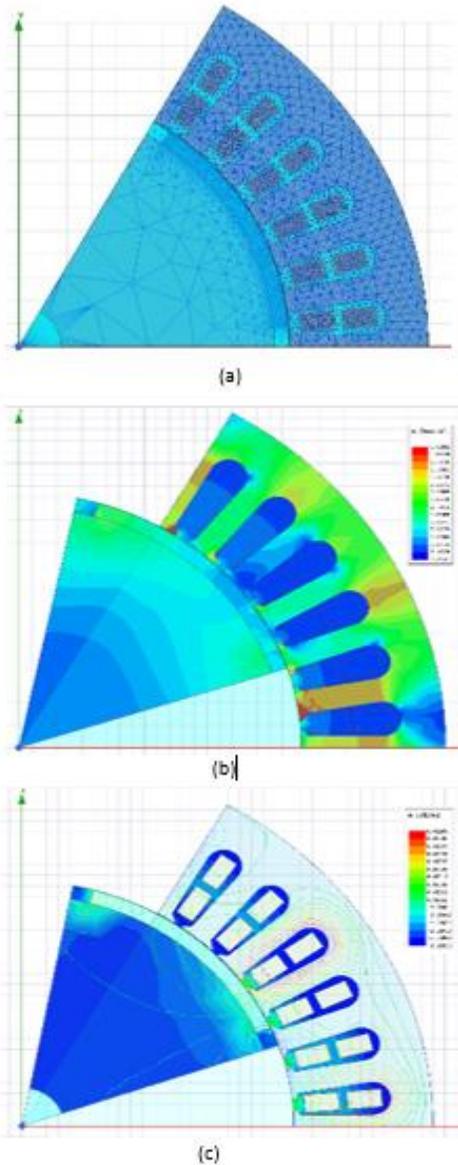


Fig.2. FEM analysis results of PMSM: a. Mesh structure, b. Magnetic flux density distribution, c. Magnetic flux lines distribution

As can be seen in Fig. 2.a, tetrahedron mesh structure has been used for the FEM based simulation model of the motor. It can be observed from Fig. 2.b indicating the magnetic flux density distribution that 1.368 T is the maximum magnetic flux density at which the core is unsaturated. Whereas it can be observed in Fig.2.c showing the magnetic flux line distribution that there is proper magnetic flux distribution with proper core geometry.

A. PMSM Inter-turns Fault

The external circuit shown in Fig. 3 was used to actualize the inter-turn fault in the FEM model of PMSM. As can be seen from the figure, the motor is driven by a 3 phase inverter circuit. The phase winding equivalent resistance changes in the motor equivalent circuit in case of an inter-turn fault. A parallel R_{fault} resistance was connected to the motor Phase A circuit in order to model this fault in the simulation. The

PMSM inter-turn winding fault was actualized at two different levels in this design. 1st Fault level is F1 and the 2nd Fault level is F2.

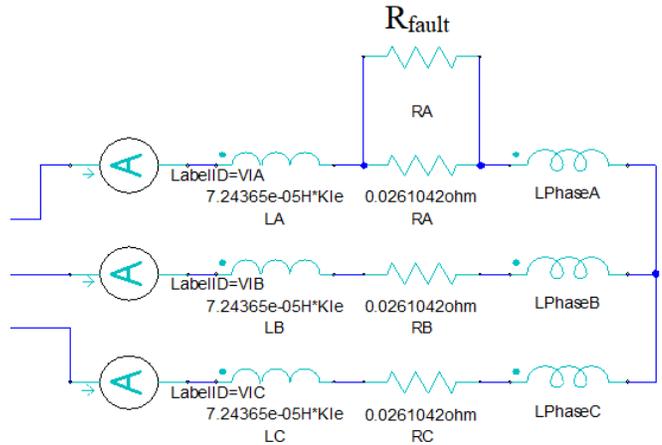


Fig.3. PMSM Inter-turns fault implementation using external circuits

III. SIGNAL ANALYSIS AND FEATURE EXTRACTION

The TMCS measured in the simulation model is a nonlinear signal. TMCS was decomposed into its IMFs in this study via EMD, one of the methods used in nonlinear signal analysis. Afterwards, the statistical method of PCC was used for determining which acquired IMF is closest to the main signal. The suggested fault detection algorithm is presented in Fig. 4.

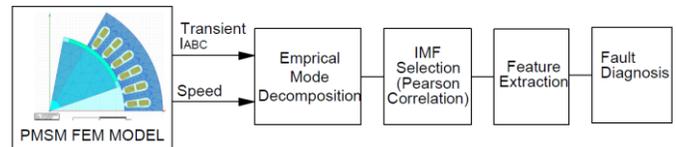


Fig.4. Suggested fault detection algorithm

A. Empirical Mode Decomposition (EMD)

It is an adaptive method developed for nonlinear and unstable signals. This method decomposes a nonlinear signal into a series of functions comprised of sub-signals known as intrinsic mode function (IMF). The EMD algorithm is shown in Fig. 5.

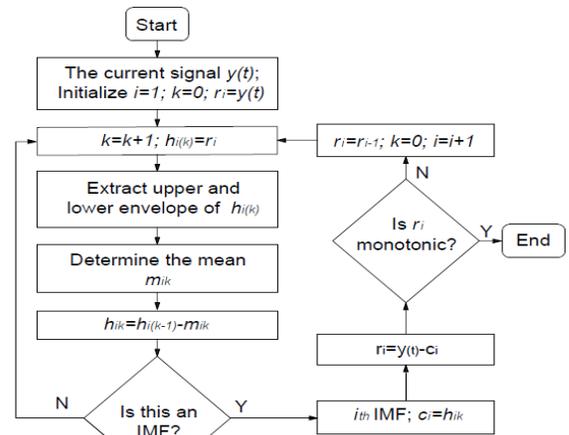


Fig.5. The EMD algorithm

An IMF function should meet the following two conditions [30]:

- (1) The number of extrema and the number of zero-crossings must either be equal or differ at most by one.
- (2) At any point, the mean value of the envelope defined by the local maxima and the envelope defined by the local minima should be zero.

A process called sifting should be carried out in order to extract each IMF. First of all, the local maxima and local minima should be determined for the input signal ($y(t)$). Afterwards, cubic spline interpolation function is applied to obtain the upper envelope with the local maxima and the lower envelope with the local minima [31]. The mean of the acquired envelopes is defined by m_1 .

$$h_1 = y(t) - m_1 \quad (6)$$

If the value of h_1 obtained as such does not meet the IMF conditions, the process steps are repeated. However, $y(t)$ is replaced by h_1 .

$$h_{1k} = h_{1(k-1)} - m_{1k} \quad (7)$$

The procedure continues until h_{1k} meets the IMF conditions. The acquired signal is the first IMF ($c_1=h_{1k}$). The step indicated below calculates the residue signal, $r_1=y(t)-c_1$ and it is controlled whether more IMF than r_1 can be calculated, that is whether it is a monotonic function or not. If r_1 is not a monotonic function, the procedure is repeated. However, this time r_1 replaces $y(t)$ in Eq. (6).

$$y(t) = \sum_{i=1}^N s_i(t) + r_i(t) \quad (8)$$

B. Pearson correlation coefficient (PCC)

After the nonlinear motor circuit signals are decomposed into its IMFs via EMD, the IMF signal with features closest to the main signal should be determined. PCC is one of the statistical methods that is most frequently used to determine the relationship between two different signals and it is also a measure of the linear correlation between two variables [32]. The covariance of the two variables in this correlation method is divided by the product of their standard deviations in this method. r , denotes the correlation coefficient and its mathematical formula is expressed by Eq. (9).

$$r = \frac{\sum x_i y_i - \frac{\sum x_i \sum y_i}{n}}{\sqrt{\left(\sum x_i^2 - \frac{(\sum x_i)^2}{n}\right)} \sqrt{\left(\sum y_i^2 - \frac{(\sum y_i)^2}{n}\right)}} \quad (9)$$

The coefficient varies between -1 and 1. Coefficients that are close to -1 indicate a negative relationship between the variables, whereas coefficients that are close to 1 indicate a positive relationship whereas the value of 0 indicates that there is no relationship between the variables.

C. Feature extraction

Extracting the features of the signal is one of the most important stages for fault detection. Using statistical parameters such as standard deviation, variance and mean value is a practical method for the analysis of stable signals in the time dimension.

The variance of a variable represents its distance from the arithmetic mean and is expressed by Eq. (10). In this equation, σ^2 , denotes the variance, n , is the number of elements in the series, x_i is the series element and \bar{x} represents the arithmetic mean of the numbers in the series.

$$\sigma^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (10)$$

Standard deviation can be expressed as the measure of the variation or fluctuation in the data set. It is frequently denoted by the symbol σ . Standard deviation is calculated by taking the square root of the variance. The standard deviation formula is presented in Eq. (11).

$$\sigma = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2} \quad (11)$$

IV. RESULTS AND DISCUSSIONS

The TMCS and motor speed were measured from the designed model for the healthy and faulty states at nominal load in order to examine the impacts of motor winding fault. Fig. 6 shows the transient currents and speed for the motor

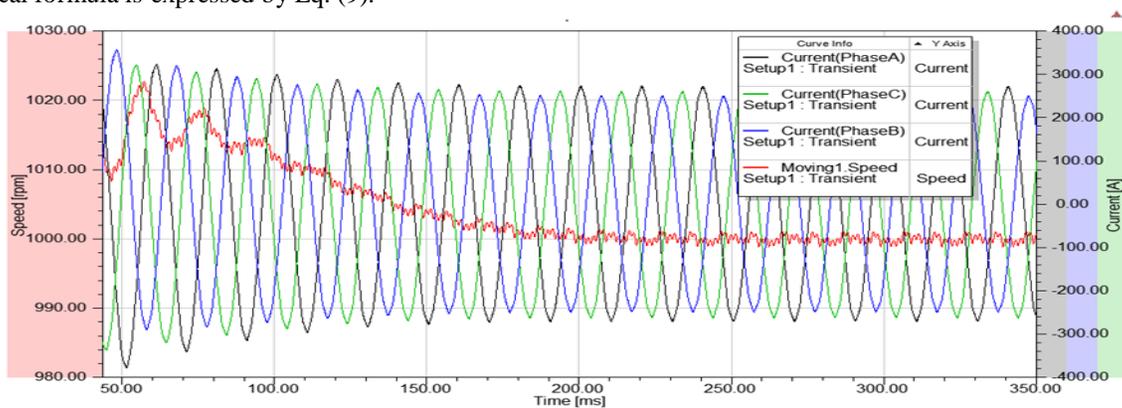


Fig.6. The TMCS and speed for the PMSM designed via FEA

The transient state that makes up the initial starting duration of the PMSM ends when the motor reaches the synchronous speed. As can be seen from Fig. 6, the motor reaches the

synchronous speed of 1000 rpm at 310 ms and the motor circuits become stable. The comparative graphs for the healthy and faulty state TMCS of the motor are presented in Fig. 7.

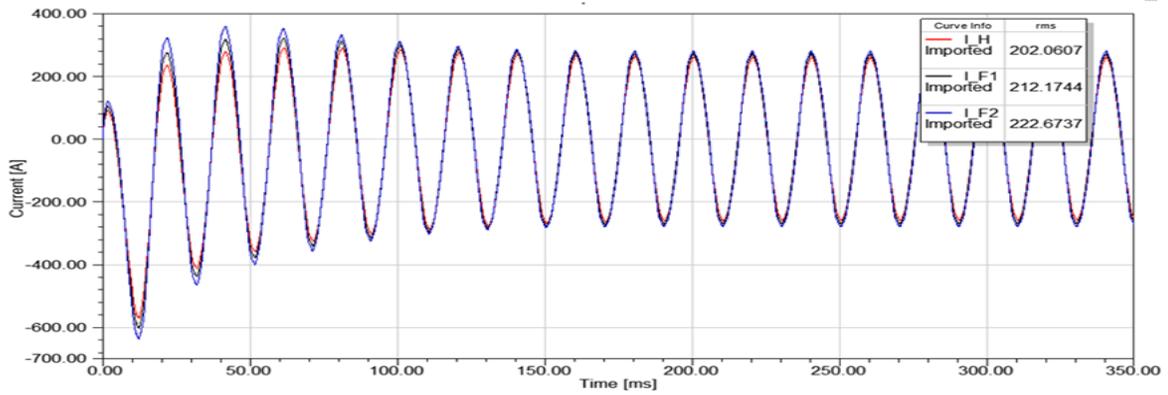


Fig.7. The TMCS for the healthy and F1, F2 faulty states of the motor

As can be seen from Fig. 7, the rms values of the motor TMCS in healthy, F1 and F2 states are 202.06 A, 212.17 A and 222.67 A respectively. The currents drawn by the motor increases with increasing fault level. Fig. 8 presents the IMFs

obtained by decomposing via EMD the nonlinear motor TMCS for the healthy, F1 and F2 states of the motor shown in Fig.7.

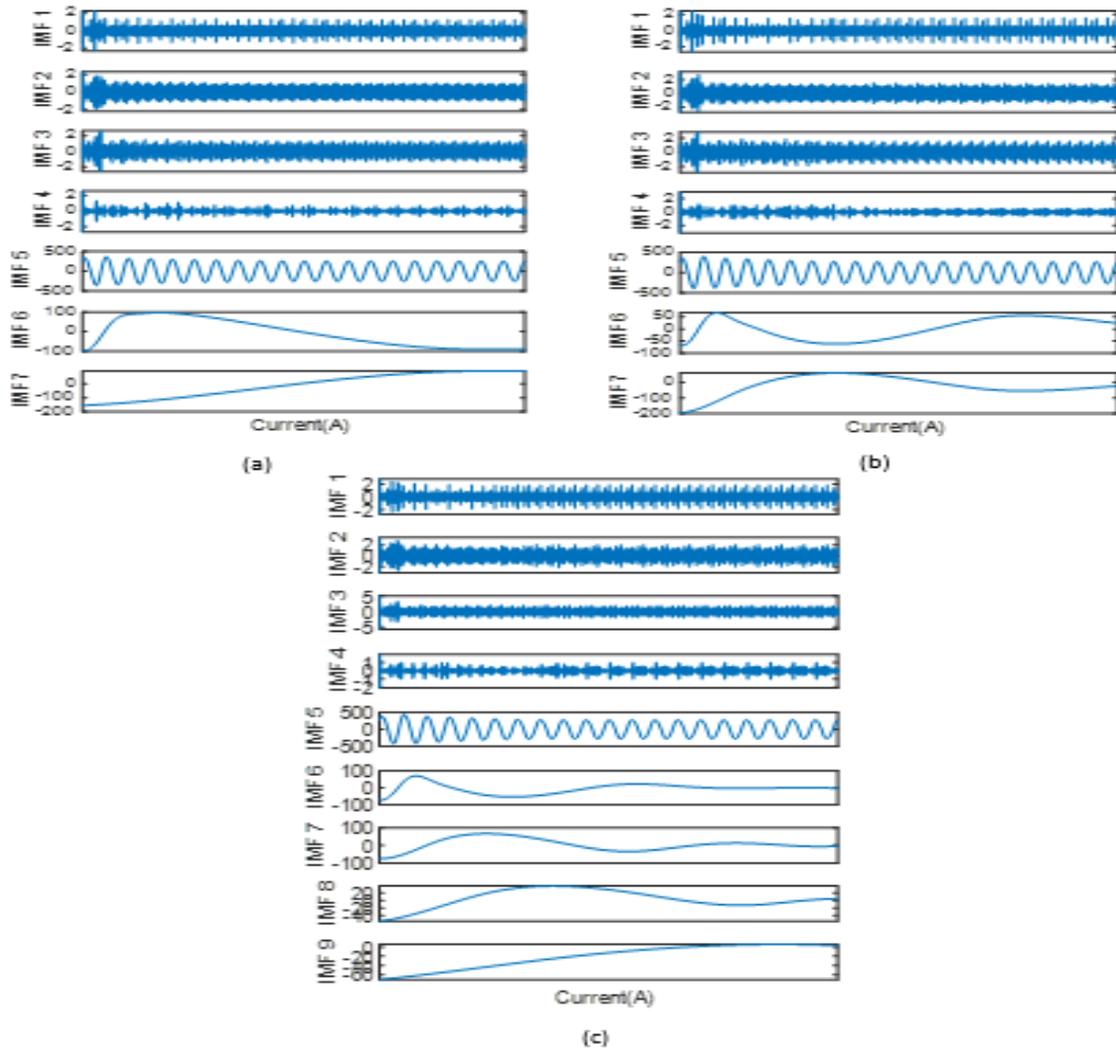


Fig.8. The EMDs of TMCSs: a). Healthy state IMFs, b). F1 fault state IMFs, c). F2 fault state IMFs.

According to the graphs in Fig. 8, the residual signal has been reached at the EMD 6th iteration (IMF6) for the healthy and F1 fault states, whereas the residual signal has been reached at the 7th iteration (IMF7) for the F2 fault state. It is difficult to visually determine the IMF representing the main signals for all three states. Hence, these IMFs were analyzed via PCC method for determining the relationship between the main signals and the IMFs and the results shown in Fig. 9 have been obtained.

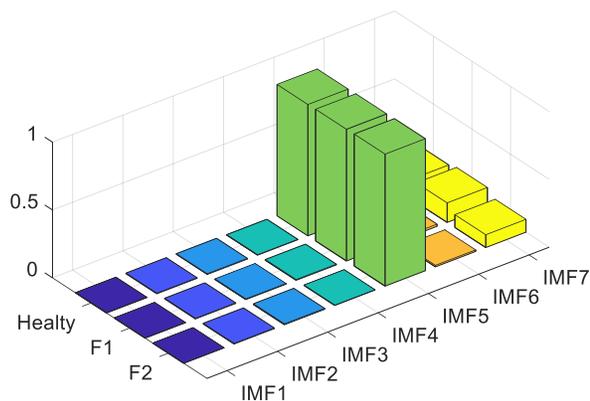
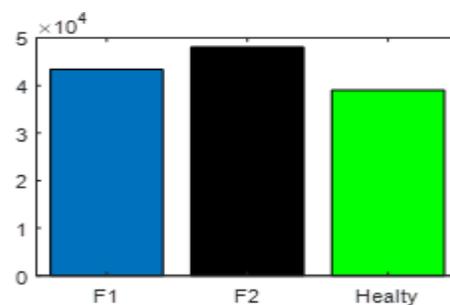


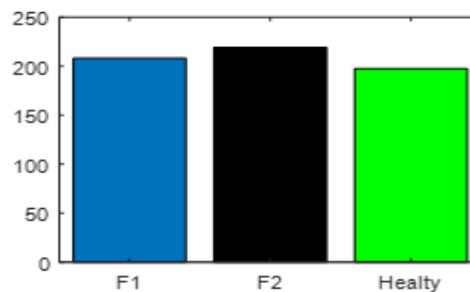
Fig.9. Correlation coefficients for the IMFs calculated for the healthy, F1 and F2 states

As an approach, it can be considered in EMD decompositions that the IMF1 obtained at the end of the 1st iteration is closest to the main signal. However, this approach may not always be correct. As can be seen from Fig. 9, since the correlation coefficients for IMF1, IMF2, IMF3, IMF4 and IMF6 are in the 0.0016684-0.007135 interval for all three states of healthy, F1 and F2; these IMFs have almost no relationship with the main signal. The correlation coefficients for IMF5 are 0.96834, 0.97077 and 0.97304 in the healthy, F1 and F2 states respectively. As can be seen, IMF5 is the signal that is closest to the main signal for all three states. Thus, information can be obtained on the motor state carried by the nonlinear main signals by way of statistical analyses of the IMF5s.

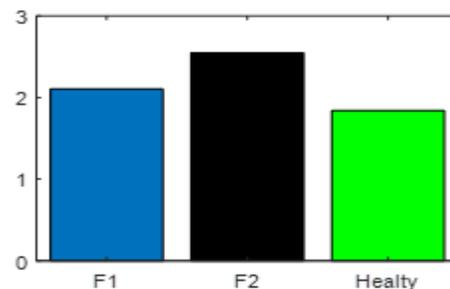
Fig. 10 presents the histograms for the statistical parameters of the IMF5 signals representing the healthy, F1 and F2 states. Whereas the variance value observed in this figure is 39034,2 for the healthy state, it is 43334.1 for the F1 state and 48049.1 for the F2 state. This value has increased from the healthy state with the increase in fault level. These changes are similar for the standard deviation and mean value parameters. The motor inter-turn winding fault and fault level can be easily determined by taking these changes into consideration.



(a)



(b)



(c)

Fig.10. Statistical parameters of the IMF5: a) Variance histogram, b). Standard deviation histogram, c). Mean value histogram

V. CONCLUSIONS

Early fault detection in motors should be carried out in temporary states during the initial starting period. In this study, a new approach based on the signal processing method of EMD and statistical analysis for winding fault detection in PMSMs using temporary state phase current.

Healthy and faulty states PMSM models were generated by using FEM. TMCS and motor speed signals were acquired by way of the designed models. The acquired signals were decomposed into their IMFs via the EMD method. At the end of the decomposition, 7 IMFs were calculated for the healthy state and F1 fault state and 9 IMFs for the F2 fault state. PCC was used for determining the IMF from among the decomposed IMFs that can represent the main signal with TMCS. It was observed for all three of the healthy and faulty states that IMF5 overlaps with the main signal with correlation coefficients of around 0.9. Fault detection and fault level detection were carried out successfully by way of statistical analyses on IMF5 signals in the healthy and faulty states. The acquired results indicate that the suggested method can be preferred for early detection of inter-turn winding faults in electrical machines.

REFERENCES

- [1] S. S. Moosavi, A. Djerdir, Y. A. Amirat, and D. A. Khaburi, "Demagnetization fault diagnosis in permanent magnet synchronous motors: A review of the state-of-the-art," *J. Magn. Magn. Mater.*, vol. 391, pp. 203–212, 2015.
- [2] H. Li, J. Hang, J. Fang, P. Zhang, S. Ding, and Q. Wang, "Inter-turn fault diagnosis of permanent magnet synchronous machine based on variational mode decomposition," *Proc. 13th IEEE Conf. Ind. Electron. Appl. ICIEA 2018*, pp. 2422–2425, 2018.
- [3] J. C. Urresty, J. R. Riba, and L. Romeral, "Diagnosis of interturn faults in pmsms operating under nonstationary conditions by applying order tracking filtering," *IEEE Trans. Power Electron.*, vol. 28, no. 1, pp. 507–515, 2013.
- [4] H. Lee, H. Jeong, and S. W. Kim, "Diagnosis of Interturn Short-Circuit Fault in PMSM by Residual Voltage Analysis," *SPEEDAM 2018 - Proc. Int. Symp. Power Electron. Electr. Drives, Autom. Motion*, pp. 160–164, 2018.
- [5] L. ERGENE and C. EKİN, "Cross Saturation Inductance Analysis of a Permanent Magnet Synchronous Motor," *Balk. J. Electr. Comput. Eng.*, vol. 6, no. 3, pp. 8–13, 2018.
- [6] Y. B. Yakut, S. Sünter, and M. Özdemir, "Simulation of Matrix Converter-Fed Permanent Magnet Synchronous Motor With Neural Fuzzy Controller," vol. 6, no. 2, 2016.
- [7] J. Härsjö, *Modeling and Analysis of PMSM with Turn-To-Turn Fault*. 2016.
- [8] S. Moon, H. Jeong, H. Lee, and S. W. Kim, "Interturn short fault diagnosis in a PMSM by voltage and current residual analysis with the faulty winding model," *IEEE Trans. Energy Convers.*, vol. 33, no. 1, pp. 190–198, 2018.
- [9] H. Jeong, S. Moon, J. Lee, and S. W. Kim, "Inter-turn short fault diagnosis of permanent magnet synchronous machines using negative sequence components," *Proceedings of the IEEE International Conference on Industrial Technology*, pp. 170–174, 2016.
- [10] H. Liang, Y. Chen, S. Liang, and C. Wang, "Fault Detection of Stator Inter-Turn Short-Circuit in PMSM on Stator Current and Vibration Signal," *Applied Sciences*, vol. 8, no. 9, p. 1677, 2018.
- [11] Y. Maanani and A. Menacer, "Modeling and Diagnosis of the Inter-Turn Short Circuit Fault for the Sensorless Input-Output Linearization Control of the PMSM," *Period. Polytech. Electr. Eng. Comput. Sci.*, vol. 63, no. 3, pp. 159–168, 2019.
- [12] A. Mohammed, J. I. Melecio, and Š. Djurović, "Open-Circuit Fault Detection in Stranded PMSM Windings Using Embedded FBG Thermal Sensors," *IEEE Sens. J.*, vol. 19, no. 9, pp. 3358–3367, 2019.
- [13] G. C. Stone, "Condition monitoring and diagnostics of motor and stator windings - A review," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 20, no. 6, pp. 2073–2080, 2013.
- [14] T. J. Kang, J. Hong, S. Bin Lee, Y. W. Yoon, D. H. Hwang, and D. Kang, "The influence of the rotor on surge pd testing of low voltage AC motor stator windings," *IEEE Trans. Dielectr. Electr. Insul.*, vol. 20, no. 3, pp. 762–769, 2013.
- [15] C. Chuang, Z. Wei, W. Zhifu, and L. Zhi, "The Diagnosis Method of Stator Winding Faults in PMSMs Based on SOM Neural Networks," *Energy Procedia*, vol. 105, pp. 2295–2301, 2017.
- [16] L. Otava, "Implementation of PMSM Inter-turn Short Fault Detection Using Frequency Analysis of Stator Currents," *IFAC-PapersOnLine*, vol. 49, no. 25, pp. 86–91, 2016.
- [17] J. C. Urresty, J. R. Riba, and L. Romeral, "Application of the zero-sequence voltage component to detect stator winding inter-turn faults in PMSMs," *Electr. Power Syst. Res.*, vol. 89, pp. 38–44, 2012.
- [18] M. Fitouri, Y. Bensalem, and M. N. Abdelkrim, "Modeling and detection of the short-circuit fault in PMSM using Finite Element Analysis," *IFAC-PapersOnLine*, vol. 49, no. 12, pp. 1418–1423, 2016.
- [19] S. Liang, Y. Chen, H. Liang, and X. Li, "Sparse representation and SVM diagnosis method inter-turn short-circuit fault in PMSM," *Appl. Sci.*, vol. 9, no. 2, 2019.
- [20] T. Yang, H. Pen, Z. Wang, and C. S. Chang, "Feature Knowledge Based Fault Detection of Induction Motors Through the Analysis of Stator Current Data," *IEEE Trans. Instrum. Meas.*, vol. 65, no. 3, pp. 549–558, 2016.
- [21] Y. Chen, S. Liang, W. Li, H. Liang, and C. Wang, "Faults and diagnosis methods of permanent magnet synchronous motors: A review," *Appl. Sci.*, vol. 9, no. 10, 2019.
- [22] E. G. Strangas, S. Aviyente, and S. S. H. Zaidi, "Time-frequency analysis for efficient fault diagnosis and failure prognosis for interior permanent-magnet AC motors," *IEEE Trans. Ind. Electron.*, vol. 55, no. 12, pp. 4191–4199, 2008.
- [23] S. Ahsanullah, Kazi; Jeyasankar, Elango; Panda, S. K.; Shanmukha, Ramakrishna; Nadarajan, "Detection and Analysis of Winding and Demagnetization Faults in PMSM based Marine Propulsion Motors" *2017 IEEE International Electric Machines and Drives Conference (IEMDC)*, pp. 1–7, 2017.
- [24] Y. Yang, J. Cheng, and K. Zhang, "An ensemble local means decomposition method and its application to local rub-impact fault diagnosis of the rotor systems," *Meas. J. Int. Meas. Confed.*, vol. 45, no. 3, pp. 561–570, 2012.
- [25] J. Rosero, L. Romeral, J. A. Ortega, and E. Rosero, "Short circuit fault detection in PMSM by means of empirical mode decomposition (EMD) and wigner ville distribution (WVD)," *Conf. Proc. - IEEE Appl. Power Electron. Conf. Expo. - APEC*, pp. 98–103, 2008.
- [26] A. Mejia-Barron, M. Valtierra-Rodriguez, D. Granados-Lieberman, J. C. Olivares-Galvan, and R. Escarela-Perez, "The application of EMD-based methods for diagnosis of winding faults in a transformer using transient and steady state currents," *Meas. J. Int. Meas. Confed.*, vol. 117, pp. 371–379, 2018.
- [27] V. T. Tran, R. Cattley, A. Ball, B. Liang, and S. Iwnicki, "Fault diagnosis of induction motor based on a novel intelligent framework and transient current signals," *Chem. Eng. Trans.*, vol. 33, pp. 691–696, 2013.
- [28] J. A. Antonino-Daviu, A. Quijano-Lopez, V. Fuster-Roig, and C. Nevot, "Case stories of induction motors fault diagnosis based on current analysis," *Pet. Chem. Ind. Conf. Eur. Conf. Proceedings, PCIC Eur.*, pp. 1–9, 2016.
- [29] H. Douglas and P. Pillay, "The impact of wavelet selection on transient motor currefnt signature analysis," *2005 IEEE Int. Conf. Electr. Mach. Drives*, pp. 80–85, 2005.

- [30] .M, Tezcan. İ, Çanakoğlu., "Asenkron Motorlarda Kırık Rotor Barı Arızalarının Sonlu Elemanlar Yöntemi ile İncelenmesi", IATS' 09, Uluslararası İleri Teknolojiler Sempozyumu, 2009. (in Turkish).
- [31] A. T. Çelebi and S. Ertürk, "Sonar imgelerinde ampirik kip ayrışımı ve morfolojik işlemler kullanarak hedef tespiti," 2010 IEEE 18th Signal Processing and Communications Applications Conference, pp. 760–763, 2010. (in Turkish).
- [32] O. Sayli, "Hilbert-huang dönüşümü ile solunum seslerindeki üfürümün saptanması," 2014 22nd Signal Process. Commun. Appl. Conf. SIU 2014 - Proc., no. April, pp. 2194–2197, 2014.(in Turkish).
- [33] H. Zhou, Z. Deng, Y. Xia, and M. Fu, "A new sampling method in particle filter based on Pearson correlation coefficient," Neurocomputing, vol. 216, pp. 208–215, 2016.

BIOGRAPHIES



Rumeysa SELCUK was born in Sivas, Turkey, in 1993. She received his B.S degree from Electrical and Electronics Department of Engineering Faculty of University of Cumhuriyet, Sivas, in 2015. She has been currently a graduate student in Electrical and Electronics Engineering Department of Institute of Science and Technology of

Gaziosmanpasa University, Tokat. She has been working as a Research Assistant in Department of Computer Engineering, Beykoz University Faculty of Engineering and Architecture.



Zafer DOGAN was born in Tokat, Turkey, in 1974. He received the B.S. and Mc. S. degree in electrical education from Marmara University, Technical Education Faculty, Istanbul Turkey, in 1996 and 2009, respectively. He is currently a Assistant Professor in the Department of Electrical And Electronics Engineering, Tokat Gaziosmanpasa University Faculty of

Engineering and Natural Sciences. His research interests are design of electrical machines and fault diagnosis in electrical machines.

Medicinal and Aromatic Plants Identification Using Machine Learning Methods

G. KAYHAN, E. ERGÜN

Abstract— In this study, different machine learning (ML) methods were used to classify medicinal and aromatic plants (MAP) namely St. John's wort (*Hypericum perforatum* L.), Melissa (*Melissa officinalis* L.), Echinacea (*Echinacea purpurea* L.), Thyme (*Thymus* sp.) and Mint (*Mentha angustifolia* L.) based on leaf shape, gray and fractal features. Naive Bayes Classifier (NBC), Classification and Regression Tree (CART), K-Nearest Neighbor (KNN), and Probabilistic Neural Network (PNN) classification were used as methods. The results indicated that plant species were successfully recognized the average of correct classification rate. The best classification rate on the NBC was taken: training data for classification rate 98.39% and test data classification rate for 98.00% are obtained. ML could be accurate tools for MAP classification tasks.

Index Terms— classification, feature extraction, image processing, machine learning.

I. INTRODUCTION

THE AUTOMATIC recognition of plants using computer vision has been commonly utilized in the literature. For example, Kalyoncu et al. used geometric features [1]. Shao et al. have classified the image sets containing the images of the leaves of the same class in each set. [2]. Ruberto et al.'s leaf classification procedure utilizes a high dimensional new features set that consist of the shape, color and texture features [3]. Du et al. [4] used fractal dimension features to recognize of plant leaf image. Xu et al. proposed a Probabilistic Neural Network (PNN) based plant classification [5]. Miao et al. suggested an evidence-theory-based rose classification method, which used many features of roses [6]. Gu et al. recommended the leaf classification method which consists of skeleton segmentation and wavelet transform and Gaussian interpolation [7]. Wang et al.'s leaf recognition methods combined shape features and a hypersphere classifier [8].

GÖKHAN KAYHAN, is with Department of Computer Engineering University of Ondokuz Mayıs University, Samsun, Turkey, (e-mail: gakayhan@omu.edu.tr).

 <https://orcid.org/0000-0003-3391-0097>

ERHAN ERGÜN, is with Department of Computer Engineering University of Ondokuz Mayıs University, Samsun, Turkey, (e-mail: erergun@omu.edu.tr).

 <https://orcid.org/0000-0003-1446-2428>

Manuscript received November 26, 2019; accepted January 08, 2020.
DOI: [10.17694/bajece.651286](https://doi.org/10.17694/bajece.651286)

Medicinal and aromatic plants (MAP) harvesting is as old as human history [9]. It is reported that at present, over 60% of the world's population, 80% in developing countries, take advantage of using plants for their medical purposes. In some Asian and African countries, alternative medicine is preferred instead of traditional medicine for healthcare needs. This is due to modern medical inadequacies and cultural preferences. [10]. The active components of the plants (*Melissa officinalis* L. [11], *Mentha piperita* L. [12] and thyme [13]) used for this purpose are concentrated in their leaves. These components are essential oils. Therefore leaves are the most important part of these medicinal and aromatic plants [14]. Studies on plant identification in the literature are widely used databases: the ImageCLEF dataset in 2011 [15] and in 2012 [16, 17], Flavia dataset [18], Leafsnap [19], Swedish leaf dataset [20], Foliage dataset [21], LifeCLEF 2015 [22]. Foliage dataset, was prepared to test leaves with various colors and patterns, contains 60 kinds of leaves [21]. The Flavia dataset contains 1907 scanned images of 32 species [18]. The Swedish leaf dataset comprises 1125 images of 15 kinds of leaves [20]. ImageCLEF dataset in 2011's [15] scanned category includes totally 3070 images and in 2012's [16] scan dataset comprises totally 6630 images. Leafsnap currently covers all 185 tree species from the Northeastern United States [19]. Odabas et al. [23] set up a MAP database, contains 752 scanned images belonging to 5 kinds of medical and aromatic plants. Main objective of this work was to classification of MAP species using computer vision for precision agriculture. MAP leaf classification is original to this work. Firstly, features extractions based digital image processing is taken and secondly ML-based classification was performed. Generally, digital morphological features [1, 18, 6, 7, 8], fractal features [4], and gray/color properties [6] commonly used. So, these powerful feature extraction methods are combined in this paper. Ruberto et al. were used 138 features in the study [3], but we use only 10 features in this study. As plant classifier methods, different ML methods (NBC, CART, PNN, and kNN classification) are used.

II. MATERIALS AND METHODS

A. Image acquisition and pre-processing

It should create the database that is needed to recognize the plant species. Obtaining appropriate images of the selected plants to create this database and extracting the features necessary for learning the machine from these images is one of the important steps.

1) Acquisition of plant leaf

Color digital images (600 dpi) were taken using a scanner (HP ScanJet G4010) of individual plants of St. John's wort (*Hypericum perforatum* L.), Melissa (*Melissa officinalis* L.), Echinacea (*Echinacea purpurea* L.), Thyme (*Thymus* sp.) and Mint (*Mentha angustifolia* L.). These images are taken from MAP database of 752 images [23]. Leaf samples were randomly taken from in experiment field populations of the MAP on grown season in 2013-2014. In this period, for catching the different phases of leaf morphogenesis, 200 St. John's Wort leaves, 200 Melissa leaves, 88 Echinacea leaves, 112 Mint leaves, and 152 Thyme leaves were collected.

2) Image pre-processing

Firstly, leaf images were converted binary images. Fig. 1 shows an example image (a), a normalized image (b), threshold binary image (c) and masked-cropped leaf image (d).

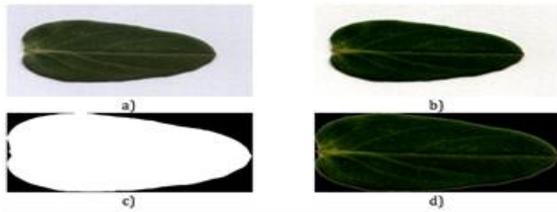


Fig.1. An example image a) raw image, b) normalized image, c) threshold binary image, and d) masked-cropped image

B. Feature Extraction

In this study, the features are consisted of the shape (3 features), gray (5 features) and fractal (2 features) properties of leaf. The total number of features is 10.

1) Shape Features (SF)

The major/minor axis length feature is a scalar that indicates the length of the major/minor axis of the ellipse (ALMaj and ALMin). Fig. 2 shows that major axis length and minor axis descriptors for binary leaf image.

Firstly, the aspect ratio representing the ratio for the ellipse with minor and major axis length, is calculated by

$$SF_{Ar} = \frac{AL_{Min}}{AL_{Maj}} \quad (1)$$

$$SF_{Min} = AR_{Min} \quad (2)$$

Eccentricity is the measure of how much any conic section deviates from being circular.

$$SF_{Sol} = \frac{A_{ROI}}{A_C} \quad (3)$$

Solidity is the measurement of the overall concavity of a particle.

$$SF_{Ecc} = \frac{E_A}{AL_{Maj}} \quad (4)$$

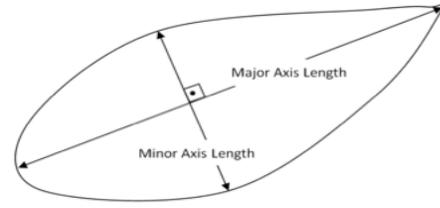


Fig.2. Ellipse's major axis length and minor axis length features for binary leaf image

2) Gray Features (GF)

The gray-level co-occurrence matrix (GLCM) is a descriptor that is defined over an image to be the distribution of co-occurring values. Contrast, correlation, energy, and homogeneity descriptors are classically calculated from GLCM. Contrast descriptor extracted the local variations from the GLCM.

$$GF_{Con} = \sum_{i,j} |i - j|^2 p(i,j) \quad (5)$$

Correlation statistic describes the joint probability occurrence of the pixel pairs.

$$GF_{Cor} = \sum_{i,j} \frac{(i - \mu_i)(j - \mu_j)p(i,j)}{\sigma_i \sigma_j} \quad (6)$$

Energy feature measures the sum of squared elements in the GLCM.

$$GF_{Ene} = \sum_{i,j} p(i,j)^2 \quad (7)$$

Homogeneity descriptor reveals the closeness of the distribution of elements in the GLCM to the GLCM diagonal.

$$GF_{Hom} = \sum_{i,j} \frac{p(i,j)}{1 + |i - j|} \quad (8)$$

And another important property is Entropy, which characterizes uncertainty of the gray leaf image.

$$GF_{Ent} = - \sum_{i,j} p(i,j) \log_2 p(i,j) \quad (9)$$

3) Fractal Features (FF)

A fractal feature of an image is an important descriptor to measure self-similar pattern that exhibits a repeating pattern that displays at every scale. The box-counting method (BCM), also known as Minkowski dimension, measures the fractal dimension of an image S in a Euclidean space R_n . In order to calculate the dimension, one should imagine a fractal S lying on an evenly-spaced grid and count the number of boxes that are necessary to cover the set. We make the grid finer by applying a box-counting algorithm to see how this number changes and the box-counting dimension is calculated using

this number. So, the local dimension distribution is calculated as,

$$df(m) = \frac{\lim_{h \rightarrow 0} \frac{N(x+h) - N(x)}{h}}{\lim_{h \rightarrow 0} \frac{r(x+h) - r(x)}{h}} \quad (10)$$

where $N(c)$ is the number of boxes of side length c required to cover the image and r is box size. In this study, we select the expected value (FF_{Mea}) of and a measure of dispersion (FF_{Std}) of df

$$FF_{Mea} = \frac{1}{M} + \sum_{m=1}^M (df(m) - FF_{Mea})^2 \quad (11)$$

$$FF_{Std} = \sqrt{\frac{1}{M} \sum_{m=1}^M (df(m) - FF_{Mea})^2} \quad (12)$$

C. ML Methods

The aim of the machine learning is that of letting huge amounts of data dictate algorithms and solutions for complex problems. At the same time, machine learning is to be contrasted with the traditional approach of depend on domain experts for the development of solutions. In this research, NBC, CART, KNN, and PNN machine learning methods were used.

1) NBC

Naive Bayes Classifiers are simple probabilistic classifiers for pattern recognition problems. This classifiers use Bayes theorem. The features which are used Bayes theorem are naive independence assumptions. Naive Bayes classifiers designed with independent predictors for each class are used, when the size of the data is high. Not much training data is needed to figure out the parameters. [24].

Bayes rule :

$p(x|C_j)$: the probability that x is a sample in class j

$P(C_j)$: first probability of class j

$p(x)$: probability of being x for any sample.

$P(C_j|x)$: The probability that sample x is from class j

$$P(C_j|x) = \frac{P(x|C_j)P(C_j)}{P(x)} \quad (13)$$

Since $p(x)$ in Equation (13) is the same for all classes, probability can be calculated as given by Equation (14). The maximum value is determined in these possibilities.

$$P(C_j|x) = P(x|C_j)P(C_j) \quad (14)$$

If all properties in Equation (14) are independent, Equation (14) is defined as Equation (15). Thus, the complexity of calculation becomes greatly reduced.

$$P(C_j|x) = \prod_{k=1}^n P(x_k|C_j)P(C_j) \quad (15)$$

$$P(C_j|x) = P(x_1|C_j)P(C_j) \times \dots \times P(x_n|C_j)P(C_j)$$

2) CART

Classification and Regression trees (CART) is a non-parametric rule-based machine learning method [25]. The main objective of CART is to separate the main data matrix (independent variables matrix) into homogeneous subgroups according to the dependent variable. The subgroups are in the form of branches of a tree and they are in a hierarchical order. In the classification tree method, Gini impurity is used to decide the purity of the binary dependent variable. In Gini algorithm, attribute values are divided into two parts.

For each partition, G_{left} (Equation 16) and G_{right} (Equation 17) are calculated.

$$G_{left} = 1 - \sum_{i=1}^k \left(\frac{|T_{class_i}|}{|B_{left}|} \right)^2 \quad (16)$$

$$G_{right} = 1 - \sum_{j=1}^k \left(\frac{|T_{class_j}|}{|B_{right}|} \right)^2 \quad (17)$$

T_{class_i} : each class value in the left part

T_{class_j} : each class value in the right part

B_{left} : all number of value in the left part

B_{right} : all number of value in the right part

Gini values (G_m) are calculated for each features by using Equation (18). The classification tree is created by selecting the feature with the smallest value from the values found.

$$G_m = \frac{1}{n} (|B_{left}| \cdot G_{left} + |B_{right}| \cdot G_{right}) \quad (18)$$

n : the total class value that belongs to a feature.

m : the number of feature

There are no classes in the regression trees. Therefore, the classification separation rules in the regression tree technique cannot be applied using the Gini index. Differences in the regression tree are performed according to the residual sum of squares. The residual sum of squares is similar to the Gini separation rules [25]. Regression trees are effective in cases of complex and nonlinear interaction with many variables [26].

3) KNN

One of the non-parametric methods used for classification and regression is the K nearest neighbors' algorithm (KNN). The algorithm determines the K nearest neighbors to the training samples of the inputs. The data is considered a member of that class if the class is dominant. K is normally a positive and small integer [27]. In Fig. 3, when the KNN algorithm is

applied (K = 3), the new data belongs to class 2 which is dominant.

KNN Algorithm:

- i. Number of neighborhoods (K) is determined.
- ii. The Euclidean distances are calculated between the input data and all data.
- iii. According to calculated distances, the data is sorted from small to large and first K data is determined.
- iv. The class information of K nearest neighbors is determined.
- v. The class of the estimated data is defined as the most repetitive class.

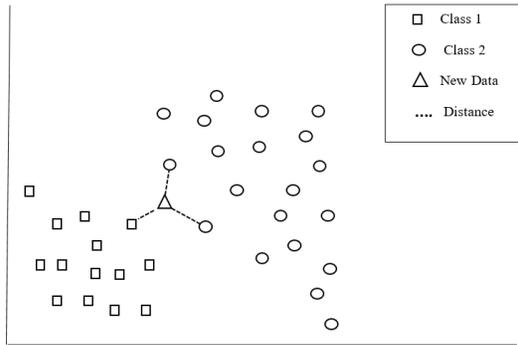


Fig. 3. Class selection for new data in the KNN algorithm

4) PNN

A Probabilistic Neural Network (PNN) is a feed forward neural network. This network was derived from Kernel Fisher discriminant analysis and the Bayesian network. A PNN network consists of four layers. this layers named the Input layer, the Pattern layer, the Summation layer, and the Output layer (Fig. 4) [28].

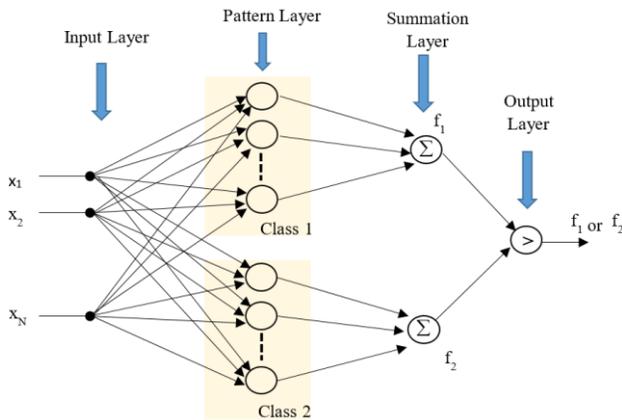


Fig. 4. A Probabilistic Neural Network (PNN)

$$g_1(x) = \frac{1}{\sqrt{(2\pi\sigma^2)^N}} e^{-\frac{\|x-\alpha^{(p)}\|^2}{2\sigma^2}} \quad (19)$$

$$g_2(x) = \frac{1}{\sqrt{(2\pi\sigma^2)^N}} e^{-\frac{\|x-b^{(q)}\|^2}{2\sigma^2}} \quad (20)$$

N: the dimension of the vectors

p: the number of vector

$\alpha^{(p)}$: p x N feature vector in class 1.

$b^{(q)}$: p x N feature vector in class 2.

x: input feature vector

σ_1 and σ_2 :spread parameters

N-dimensional data is passed through the gauss functions that belongs to class 1 and class 2 in the pattern layer (Equation 19 and 20).

The sum of gauss outputs of each class are f1 and f2 respectively.

If $f_1 > f_2$ then x belongs to class 1.

D. Evaluation Strategy

1) k-fold Cross Validation:

The k-fold cross-validation method is widely used to evaluate objectively the classifier performance. This method, by random selection, allows all instances to enter the test set. For the M data sets, k-fold cross validation algorithm comprises the following steps,

Step1. The data set is divided into k-random fold. Thus, each fold M/k will have data.

Step 2. Each i. belonging to the fold (i = 1..k) data consists of test dataset, others consists of the train dataset. ML network was trained with each fold's train set, and then network was evaluated with the test set.

2) Accuracy:

A matrix-shaped special table is used to visualize the performance of a consulting learning algorithm in machine learning applications [29]. Each column and row of the matrix represents the instances in a predicted class and the instances in an actual class respectively. To ISO 5725-1, it is known that the general term accuracy (Equation 21) is used to describe the closeness of measurement to the true value [30]. That is, the accuracy is the proportion of both true positive and true negative results among the total number of cases tested [31].

$$Accuracy(\%) = \frac{\text{numbers of true features}}{\text{total numbers of samples}} \times 100 \quad (21)$$

III. RESULTS AND DISCUSSION

In this research, k=10 cross validation is used. So total samples are 752, train samples are 677 (90%) and test samples are 75 (10%). NBC, CART, KNN and PNN were used for identification of some medicinal and aromatic plant species. The training accuracy obtained in all classifier is higher (Table I). The highest performance is obtained by NBC in dataset (98.39%, 98.00%). This high performance is directly related to the method followed in the feature extraction.

TABLE I
PERFORMANCE RESULTS WHICH WERE OBTAINED THE DIFFERENT ML METHODS

ML Methods	Data Set	Accuracy
NBC	Training	98.39%
	Testing	98.00%
CART	Training	99.51%
	Testing	97.61%
KNN	Training	100.00%
	Testing	81.92%
PNN	Training	100.00%
	Testing	67.69%

Each feature (Gray, Shape, and Fractal features) was removed from the model and its contribution to performance was determined. The accuracy decrement when they have removed is shown in Table II.

TABLE II
CONTRIBUTION OF PROPOSED FEATURES WITH REGARDS TO ACCURACY IN MAP DATABASE.

Item	Contribution (%)
Gray features	5.97
Shape features	3.28
Fractal features	0.04

The high test accuracy (98.00 %) was obtained by NBC, for all folds, total training confusion matrix are given in Fig. 5 and total testing confusion matrix in Fig. 6.

Actual Class	%	Predicted Class				
		St. John's wort	Melissa	Echinacea	Thyme	Mint
St. John's wort	98.50	0.00	1.50	0.00	0.00	
Melissa	0.00	99.28	0.11	0.00	0.61	
Echinacea	2.53	1.14	96.34	0.00	0.00	
Thyme	0.00	0.66	0.00	99.34	0.00	
Mint	0.00	2.08	0.99	0.00	96.93	

Fig. 5. Confusion Matrix for Classification kNN for train

The 6.659 out of 6.768 train samples were correctly classified (98.39%). In Fig. 5, Echinacea type was classified incorrectly St. John's wort (2.53%) and Melissa (1.14%). Similarly, Thyme type was classified incorrectly Melissa (0.66%). Melissa was confused to mostly Mint (0.6%).

Actual Class	%	Predicted Class				
		St. John's wort	Melissa	Echinacea	Thyme	Mint
St. John's wort	98.50	0.00	1.50	0.00	0.00	
Melissa	0.00	98.50	0.50	0.00	1.00	
Echinacea	1.14	1.14	95.45	0.00	2.28	
Thyme	0.00	0.66	0.00	99.34	0.00	
Mint	0.00	2.68	0.89	0.00	96.43	

Fig.6. Confusion Matrix for Classification kNN for test

In Fig.6, 737 out of 752 test samples were correctly classified (98.00%). Echinacea type was classified incorrectly St. John's wort (1.14%) , Melissa (1.14%) and Mint (2.28%). Similarly, Thyme type was classified incorrectly Melissa (0.66%). Melissa was confused to mostly Mint (1%). Misclassification style for train dataset is similar to test dataset.

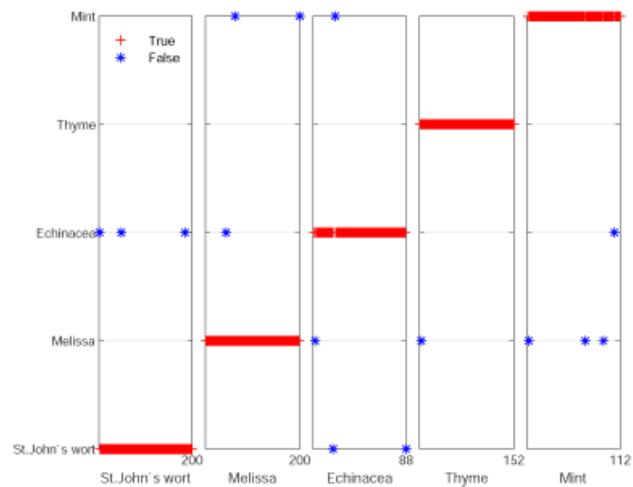


Fig. 7. Result of the classification NBC for test dataset

Wu et al. used 12 features (orthogonalized 5 principal variables) to 32 leaf types and achieved %90.31 accuracy rate with PNN classification [17]. Du et al. used 15 features to 20 leaf types and achieved %93 accuracy rate with 1-NN classification [32]. Du et al. used 20 features to 35 leaf types and achieved %87.14 accuracy rate with k-NN classification [4]. Kadir et al. used 17 features to 32 leaf types and achieved %94.07 accuracy rate with PNN classification [21]. Ruberto et al. used 138 features to 32 leaf types and achieved approximately %100 accuracy rate with SVM classification [3]. Kalyoncu et al. used a lot of features to 32 leaf types and achieved approximately %94 accuracy rate with LDC

classification [1]. In this work, it achieves %98.00 accuracy rate with NBC, which there are 5 leave types and 11 features. Fig.7 shows the correct and incorrect classification in test dataset. For instance, St. John's wort leaves, which have 200 test data, were correctly classified the numbers of 197 data (+ symbol) and incorrectly classified the numbers of 3 data (* symbol) as Echinacea leaves.

IV. CONCLUSION

Due to the difficulty of morphologically similar leaves, plant recognition based on images is difficult processing. Therefore, the performance of many algorithms used in plant identification has low accuracy.

In this study, the leaves of medicinal and aromatic plants were automatically classified according to leaf shape and color using computer vision. For this classification, a new feature extraction approach, which used gray, shape and fractal features of leaf, was proposed. The used feature set is a combination of several known feature sets. Additionally, we proposed the use of NBC that has a higher accuracy compared to KNN, CART, and PNN classifiers. In this research, it has been shown that leaf identification can be developed by combining features complementing each other. In this study, the properties that increase the accuracy of plant recognition are combined. Among the proposed methods, NBC is a better-class classifier. The experimental results on MAP database verify the effectiveness of the proposed method.

The reason is that NBC based on applying Bayes' theorem with naive independence assumptions among the gray, shape and fractal features. This allows the use of a combination of independent leave's features. Additionally, online classification time of NBC is the lowest due to single scan and fast classify. Thus, it is very pertinent for large-scale applications.

REFERENCES

- [1] C. Kalyoncu, T. Önsen: "Geometric leaf classification", *Computer Vision and Image Understanding*, 2015, 133, (0), pp. 102–109
- [2] M. Shao, J. Du, J. Wang, C. Zhai: "Recognition of leaf image set based on manifold-manifold distance", In: *Intelligent Computing Theory - 10th International Conference, ICIC 2014, Taiyuan, China, August 3-6, 2014, Proceedings*, pp. 332–337
- [3] C.D. Ruberto, L. Putzu: "A fast leaf recognition algorithm based on SVM classifier and high dimensional feature vector", In: *VISAPP 2014 - Proceedings of the 9th International Conference on Computer Vision Theory and Applications, Volume 1, Lisbon, Portugal, 5-8 January, 2014*, pp. 601–609
- [4] J. Du, C. Zhai, Q. Wang: "Recognition of plant leaf image based on fractal dimension features", *Neurocomputing*, 2013, 116, pp. 150–156
- [5] S.G. Wu, F.S. Bao, E.Y. Xu, Y. Wang, Y. Chang, Q. Xiang: "A leaf recognition algorithm for plant classification using probabilistic neural network", *CoRR abs/0707.4289*, 2007
- [6] Z. Miao, M. Gandelin, B. Yuan: "An oopr-based rose variety recognition system", *Eng. Appl. of AI*, 2006, 19, (1), pp 79–101
- [7] X. Gu, J.-X. Du, X.-F. Wang: "Leaf recognition based on the combination of wavelet transform and gaussian interpolation", In: Huang, D.-S., Zhang, X.-P., Huang, G.-B. (eds.) *Advances in Intelligent Computing, Lecture Notes in Computer Science*, Springer, 2005, vol. 3644, pp 253–262
- [8] X. Wang, J. Du, G. Zhang: "Recognition of leaf images based on shape features using a hypersphere classifier", In: *Advances in Intelligent Computing, International Conference on Intelligent Computing, ICIC 2005, Hefei, China, August 23-26, 2005, Proceedings, Part I*, pp 87–96
- [9] S.S. Dhillion, H. Svarstad, C. Amundsen, H. Bugge: "Bioprospecting: Effects on Environment and Development", *Ambio*, 31, (6), 2002, pp:491–493,
- [10] J.R.S. Tabuti, S.S. Dhillion, K.A. Lye: "Traditional medicine in bulamogi county, uganda: its practitioners, users and viability", *Journal of Ethnopharmacology*, 85, (1), 2003, pp 119–129
- [11] H. Sadraei, A. Ghannadi, K. Malekshahi: "Relaxant effect of essential oil of melissa officinalis and citral on rat ileum contractions", *Fitoterapia*, 2003, 74, (5), pp 445–452
- [12] M. Mucciarelli, W. Camusso, C.M. Berteau, S. Bossi, M. Maffei: "Effect of (+)-pulegone and other oil components of menthapiperita on cucumber respiration", *Phytochemistry*, 2001, 57, (1), pp 91–98
- [13] H. Baydar, O. Sağdıç, G. Özkan, T. Karadoğan: "Antibacterial activity and composition of essential oils from *origanum*, *thymbra* and *satureja* species with commercial importance in Turkey", *Food Control*, 2004, 15, (3), pp 169–172
- [14] M. Centritto, F. Loreto, A. Massacci, F. Pietrini, M.C. Villani, M. Zacchini: "Improved growth and water use efficiency of cherry saplings under reduced light intensity", *Ecological Research*, 2000, 15, (4), pp. 385–392
- [15] "The CLEF 2011 plant images classification task", <http://ceur-ws.org/Vol-1177/CLEF2011wn-ImageCLEF-GoeauEt2011a.pdf>, accessed 25 June 2015
- [16] "The imageclef 2012 plant identification task", <http://ceur-ws.org/Vol-1178/CLEF2012wn-ImageCLEF-GoeauEt2012.pdf>, 25 June 2015
- [17] B. Yanikoglu, E. Aptoula, C. Tirkaz: *Automatic plant identification from photographs Machine Vision and Applications*, 2014, 25, pp.1369–1383
- [18] S.G. Wu, F.S. Bao, E.Y. Xu, Y. Wang, Y. Chang, Q. Xiang: "A leaf recognition algorithm for plant classification using probabilistic neural network", *CoRR abs/0707.4289*, 2007
- [19] N. Kumar, P.N. Belhumeur, A. Biswas, D.W. Jacobs, W.J. Kress, I.C. Lopez, J.V.B. Soares: "Leafsnap: A computer vision system for automatic plant species identification", In: *Computer Vision - ECCV 2012 - 12th European Conference on Computer Vision, Florence, Italy, October 7-13, 2012, Proceedings, Part II*, pp. 502–516
- [20] O.J.O. Söderkvist: "Computer vision classification of leaves from swedish trees". Master's thesis, Linköping University, SE-581 83 Linköping, Sweden, LiTH-ISY-EX-3132, September 2011
- [21] A. Kadir, L.E. Nugroho, A. Susanto, P.I. Santosa: "Neural network application on foliage plant identification", *CoRR abs/1311.5829*, 2013
- [22] M.M. Ghazi, B. Yanikoglu, E. Aptoula: "Plant identification using deep neural networks via optimization of transfer learning parameters", *Neurocomputing*, 2017, 235, pp. 228–235
- [23] M.S. Odabas, N. Senyer, G. Kayhan, E. Ergun: "Estimation of chlorophyll concentration index at leaves using artificial neural networks", *Journal of Circuits, Systems, and Computers*, 2015
- [24] C.D. Manning, P. Raghavan, H. Schütze: "Introduction to Information Retrieval", Cambridge University Press, New York, NY, USA, 2008
- [25] L. Breiman, J.H. Friedman, R. Olshen, A.C.G. Stone: "Classification and regression trees", Wadsworth International Group, Belmont, California, USA, 1984.
- [26] W.-Y. Loh: "Classification and regression trees", *Wiley Interdisc. Rev.: Data Mining and Knowledge Discovery*, 2011, 1, (1), pp. 14–23
- [27] N.S. Altman: "An introduction to kernel and nearest-neighbour nonparametric regression", 1992, 46, (3), pp. 175–185
- [28] D.F. Specht: "Probabilistic neural networks", *Neural Netw.* 1990, 3, (1), pp. 109–118
- [29] S.V. Stehman: "Selecting and Interpreting Measures of Thematic Classification Accuracy", *Remote Sensing of Environment*, 1997, 62, (1), pp. 77–89
- [30] BS ISO 5725-1: "Accuracy (trueness and precision) of measurement methods and results - part 1: General principles and definitions", 1994
- [31] C.E. Metz: "Basic principles of {ROC} analysis", *Seminars in Nuclear Medicine*, 1978, 8, (4), pp 283–298
- [32] J.-X. Du, X.-F. Wang, G.-J. Zhang: "Leaf shape based plant species recognition", *Appl. Math. Comput.*, 2007, 185, (2), pp 883–893

BIOGRAPHIES



GÖKHAN KAYHAN Trabzon, in 1978. He received the B.S. degrees in Electrical and Electronics Engineering from the Ondokuz Mayıs University, Samsun, in 1999 and the M.S. degrees in Electrical and Electronics Engineering from the Ondokuz Mayıs University, Samsun, in 2004 the Ph.D. degree in the B.S. degrees in Electrical and Electronics Engineering from the Ondokuz Mayıs University, Samsun, in 2011.

From 2001 to 2013, he was a Research Assistant in Department of Computer Engineering at Ondokuz Mayıs University. Since 2013, he has been an Assistant Professor Dr. at the Department of Computer Engineering, Ondokuz Mayıs University. His research interests include machine learning artificial intelligence, image processing, signal processing.



ERHAN ERGÜN Samsun, in 1970. He received the B.S. degrees in Electrical and Electronics Engineering from the Hacettepe University, Ankara, in 1993; M.S. and Ph.D. degrees in Electrical and Electronics Engineering from the Ondokuz Mayıs University, Samsun, in 1999 and 2008.

From 2009 to 2010, he was a Assistant Professor Dr. at Bafra Vacation School, Ondokuz Mayıs University. Since 2012, he has been an Assistant Professor Dr. at the Department of Computer Engineering, Ondokuz Mayıs University. His research interests include analog electronic circuits, image processing, signal processing.

Hexapod Robot Design and Performance Comparison of Fuzzy and PID Control Methods

S. AYDIN, L. GOKREM and M. S. CAN

Abstract— In this study, a six-leg spider robot (hexapod) was designed and controlled for greenhouse, search and rescue operations and military applications. Solidworks design program was used in the design stage and Matlab Simulink program was used in the control stage of the robot. In the study, a motion trajectory was determined by interpolation technique for the robot and it was aimed to move on this trajectory. Proportional-Integral-Differential (PID) controller and Fuzzy Logic Controller (FLC) were used for the trajectory control of the robot. The robot walk movements were applied in both control types with motion on flat ground, motion on inclined ground, motion at different weights and different friction coefficients. In the analysis studies, the total weight of the robot was taken into consideration as 4kg. As a result of the analysis studies, it was observed that this robot, which was designed and analyzed, followed a trajectory defined by a mass of 4 kg with an error value of 1 mm on average. During the trajectory tracking, it was found that the Fuzzy controller performs better than the PID controller at the turning points of the reference trajectory curve of the robot.

Index Terms— Solidworks, Matlab, PID, fuzzy logic, Simulink, hexapod robot.

I. INTRODUCTION

THE WORD robot, which has entered many areas of our lives, was first used in history by Czech writer Karel Capek in a playwriting in 1921. In the Czech language, the word robot means captive, slave. Nowadays, the word robot is used for machines that can behave like human beings and can be used as an alternative to human beings.

Throughout history, human beings have endeavored to

SEFA AYDIN, is with Department of Mechatronic Engineering, Tokat Gaziosmanpasa University, Tokat, TURKEY
(e-mail: sefa.aydin.29@hotmail.com).

LEVENT GÖKREM, is with Department of Mechatronic Engineering, Tokat Gaziosmanpasa University, Tokat, TURKEY
(e-mail: levent.gokrem@gop.edu.tr).

 <https://orcid.org/0000-0003-2101-5378>

MEHMET SERHAT CAN, is with Department of Electrical-Electronic Engineering, Tokat Gaziosmanpasa University, Tokat, TURKEY
(e-mail: mehmetserhat.can@gop.edu.tr).

 <https://orcid.org/0000-0003-2356-9921>

Manuscript received November 25, 2019; accepted December 12, 2019.
DOI: [10.17694/bajece.650784](https://doi.org/10.17694/bajece.650784)

develop tools to help them, and tried to make animals and plants like devices, especially similar to human. Robot experiments were created with mechanical devices in different parts of history and the actual development emerged in the 20th century. In the 1900s, robots came into our lives as auxiliary work machines in weaving looms [1].

With the development of technology, the studies on autonomous control of the systems have accelerated and thus the power of production systems is changing and developing day by day. The most important factors of this development and change are undoubtedly the robots used in the systems. Autonomously controlled robots are capable of processing and receiving images without any external effects. Since they are stable and can work continuously, the quality of the results is minimized by reducing the margin of error. Robots are widely used in medicine, defense, aircraft and space studies outside the industry.

Hexapod robots include legs attached to six robot bodies. The legs are controlled to certain degree, so that the robot can move within the environment and perform the specified tasks. Such robots are suitable for terrestrial and space studies. They include versatile movement, variable geometry, good stability, and durable movement in various terrains [2,3]. Hexapod robots are the subject of many applications because they can overcome larger obstacles than wheeled robots of the same size. The use of wheels or pallets limits movement if there are obstacles that can extend up to half the diameter of the wheels. In contrast, legged robots can overcome comparable obstacles as the foot length is longer than a wheel diameter [2,4].

Nowadays, in many machine production processes, a model of the machine is made with 3D computer aided design (CAD) program and also heating, abrasion and strength tests are performed on this model with analysis programs. As a result of the analysis, the most suitable design model is obtained and with this method, rapid and less costly prototypes can be obtained. In this study, the 3D design of a six-legged hexapod (spider like) robot and its control on Matlab / Simulink base were used to control the factors such as spraying, irrigation or collecting at the right time in the greenhouse farming area without any external intervention, autonomous control, and transferring images from areas where human health might be at risk. For the movement of the robot, the points that the robot must go within the motion range of the robot are defined and a trajectory is defined for these points by using curve fitting method. FLC and PID controllers were used to monitor the trajectory of the robot and the control performances of

these two controllers were compared. The actual prototype of the robot must be equipped with cameras for observation and application-specific sensors for data collection. In this way, images taken from the camera can be used in processes such as obstacle recognition, analysis of plant leaves, mine detection and image transfer using image processing and machine learning algorithms.

II. MATERIAL

A. Simulation Softwares

SolidWorks is a three-dimensional design program that uses Microsoft Windows graphic interfaces. SolidWorks is now widely used in mechanical designs. The design phase consists of a special opened part command for each part and assembly commands that combine these parts. It begins with a two-dimensional drawing at the beginning of a design. After the main body of the part is drawn, three dimensionality is given to the drawing by means of commands. Iron and steel materials can be assigned to these designed parts. Parts can be painted in desired colors. Fittings are combined in the "assembly" section of the completed design. The relationship to the point of contact of each part is defined. These relations are concentric, tangent to each other or coincident. It is possible to make corrections and new part designs on the parts when the assembly stage is started [4].

Matlab mathematical programming language is developed by MathWorks. Matlab, which contains many ready-made commands, allows users to work with programs created in different languages such as C, C++ and Java. Matlab is now widely used in engineering.

B. PID Controller

Control with PID controller is a common control method in robotics and automation systems. This method receives continuous feedback from the system and calculates the error value between the reference point and the controlled value within the specified time and generates output values to reduce the error to zero. Fig. 1 shows a closed-loop PID controller control block diagram. In general, in a closed-loop control structure, the system output signal $y(t)$ is intended to be output to a reference value $r(t)$ in the shortest possible time, with minimal overrun. In addition, $y(t)$ should be able to hold in spite of external disturbances at this reference value. This is called durability. The PID controller takes the error value $e(t)$ as an input and finds the control signal $u(t)$ by summing the results obtained using proportional (P), integral (I) and derivative (D) operators. Finally, the control sign $u(t)$ is applied to the system. Thus, the error value $e(t)$ is intended to be minimized.

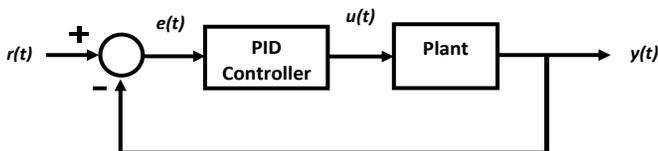


Figure 1. PID controller basic diagram.

The PID controller's performance is determined by the proportions, derivative and integral operators' gain coefficients. These gain coefficients are expressed as K_p , K_i and K_d , respectively. These gain values very important for rise time, overshoot rate, peak value, settling time and steady state error values of the system. Some energy storage systems do not respond immediately to input signals and react temporarily over time [5]. Therefore it is very important to tuning the gain coefficients of the PID controller. The expressions of the proportional, integral and derivative operators in the time domain are as shown in Equ.(1) -Equ.(4).

$$P_{operator} = K_p e(t) \quad (1)$$

$$I_{operator} = K_i \int_{t_1}^{t_2} e(t) dt \quad (2)$$

$$D_{operator} = K_d \frac{de(t)}{dt} \quad (3)$$

$$u(t) = K_p e(t) + K_i \int_{t_1}^{t_2} e(t) dt + K_d \frac{de(t)}{dt} \quad (4)$$

The expression of the control signal with PID controller in the Laplace space is given in Equ. (5) below and the transfer function of the PID controller in the Laplace space is given in Equ. (6).

$$U(s) = \left(K_p + \frac{K_i}{s} + K_d s \right) E(s) \quad (5)$$

$$C(s) = \frac{U(s)}{E(s)} = K_p + \frac{K_i}{s} + K_d s \quad (6)$$

The degrees of the I and D operators given above are expressed as integers and the representation is called integer order PID or PID. The degrees of the I and D operators are not integer degrees, λ for the integral operator and μ for the derivative operator. These two parameters are in the range of $[0,1]$ and in this case the fractional order PID controller is obtained. In the fractional order PID design, in addition to the P , I and D parameters, two additional parameters indicated by λ and μ are added, and this allowing the use of five parameters. Fractional integral and derivative operations are found in L'Hospital's work [6,7].

C. Fuzzy Logic Controller

FLCs can be said to be one of the most popular methods used to control nonlinear systems in particular. FLC is one of the model independent methods. Using the expert knowledge in control design for the strengths of FLCs, it can be said that the irregularity system can increase both its flexibility and robustness in the presence of irregularities and uncertainties. Because of these properties, FLCs are preferred in the control problems of mobile robots, especially in the control of unmanned aircraft, which, as a more challenging control

problem, previously forced the acquisition of mathematical model parameters [8-14].

In FLCs, input variables are correlated with each other on the basis of fuzzy set theory, resulting in output variables. The way in which input variables are associated and what results are obtained from these relationships are performed by rules. Rules are created within the framework of the user's knowledge and expertise. An FLC creates four basic units; fuzzification, rules, inference, defuzzification units. An FLC architecture is shown in the diagram in Fig. 2.

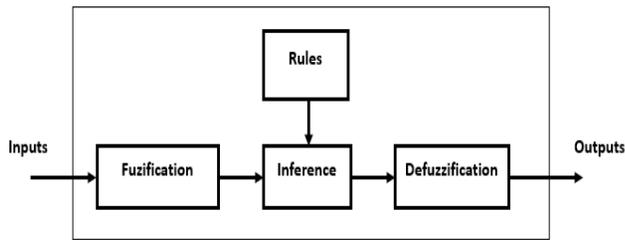


Figure 2. FLC system architecture.

Fig. 2 shows a FLC structure. In FLC, the input data is first converted to fuzzy subsets within the [0.1] range of the blur step. Membership functions are used for this conversion. First, several fuzzy subsets are defined for each input variable, including the range of change of the input variable. The instantaneous value of the input variable, whichever fuzzy subset corresponds, applies to the membership function of that subset and fuzzy values are calculated.

Another step in FLC design is the rule base. Rules are created according to the expertise of the designer. These rules are in the form of “if-then” definitions. In the inference phase, fuzzy outputs are obtained according to fuzzy input values and rules. The final step in FLC design is the rinsing step. In the rinsing step, the process of converting the fuzzy results obtained by the extraction step into a real and controllable system is performed. The most well-known methods are “Mamdani”, “Sugeno” and “Tsukamoto”.

The designed FLC is used in place of the PID controller in the closed loop control structure given in Fig. 1. The FLC controller produces more effective results than the PID controller, especially in the control of systems involving parameter changes.

III. METHOD

Solidworks design program was used in the robot design stage and Matlab R2017b / Simulink libraries were used in the motion control stage of the robot. The design is based on the spider anatomical structure and the leg structures are designed to be 60 degrees apart to maximize the balance factor. In the design, the telescopic camera carrier, which can be extended and shortened, is placed in the center of gravity of the robot for different applications. The robot is intended to move on an orbit. For this reason, an orbit is defined by curve fitting method. Thus, the robot can be sent to the desired coordinates in the x - y plane in the cartesian coordinate system in the

desired time. Since the robot moves in two dimensions in the x - y plane, a controller for trajectory control on the x -axis and a second controller for control on the y -axis are used. In the x -axis, a linear signal is referred to as the reference, and the y -axis is referred to as a path (trajectory) of determined points. The autonomous motion control of the robot on these trajectories has been done with PID and FLC controllers separately in Simulink environment. The tripod gait was used for the stepping motion of the robot. Wave stepping is a form of walking motion that is commonly used in spider robots. Fig. 3 shows tripod gait movements.

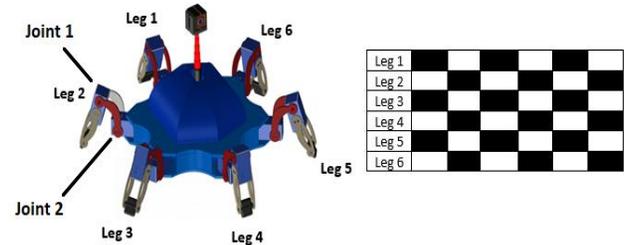


Figure 3. Hexapod robot and tripod gait motion.

As shown in Fig. 3, each leg of the robot has two joints. There are servo motors in these joints to provide movement. All of these servo motors are centered at an angle of 60 degrees during assembly. The first joints in the legs connect the foot to the body of the robot and at the same time move it back and forth in the x - y plane. The second joint performs the up and down movement of the foot in the x - z plane. In the basic robot movement, according to the table order in Fig. 3, firstly, two joints in the 1st, 3rd and 5th legs are given a sinus sign together, the joints of the 2nd, 4th and 6th legs having the same amplitude and frequency immediately after the contact of these legs to the ground sinus sign is applied. If the step angles on the 1st, 3rd and 5th legs are larger than the step angles on the 2nd, 4th and 6th legs, the robot moves in the x direction and vice versa. In the absence of trajectory control with PID and FLC controllers, the robot moves randomly with a tripod gait. The control mark of the controller controlling the movement in the x direction adjusts the pitch angles of the 1st, 3rd and 5th legs, and the control mark of the controller in the y direction adjusts the step angles of the 2nd, 4th and 6th legs. The control signal obtained from the PID and FLC controllers changes the amplitude of the sine signals providing tripod gait motion, thereby increasing or decreasing the pitch angle of the robot. When the trajectory is defined and this trajectory is given as reference to the PID and FLC controllers, the robot walks on this trajectory with wave velocity.

During the experimental simulation studies at Simulink, the weights of all parts of the robot were taken to be the same as the limits of the part weights actually planned. Experimental simulation studies were carried out in both control types and experimental values were obtained for motion on flat ground, motion on sloping ground, different weights and different friction coefficients. Based on the values obtained, FLC and PID controllers were compared.

A. Design of the hexapod robot

At this stage, parts are designed for robot formation in the dimensions, thicknesses and patterns determined in the SolidWorks design program environment. The designed robot has six legs. Inspired by spiders in nature, this six-legged chassis structure is structured in such a way that the legs are spread out in three parts to the right and left as spiders have leg and tail structure. Since the designed robot will include the camera structure and the elongated neck structure, a chassis design has been designed so that the legs are placed at 60 degree intervals in order to maximize the balance by ignoring the tail and leg portions. The preliminary drawing of the design is shown in Fig. 4.

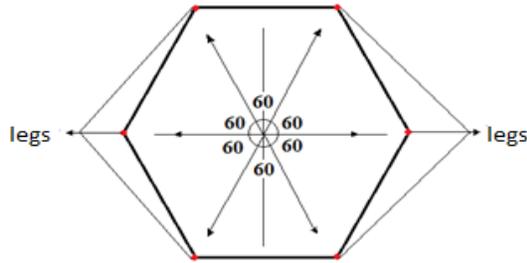


Figure 4. Chassis front drawing.

In the drawing of the frame, drawing between the legs of the robot was made to the right and left of each leg using a depth of 15 mm of a circle with a diameter of 73 mm. The plate thickness is 0.6 mm and the clearance between the two legs is designed as 94.5 mm. The process is shown in Fig. 5.

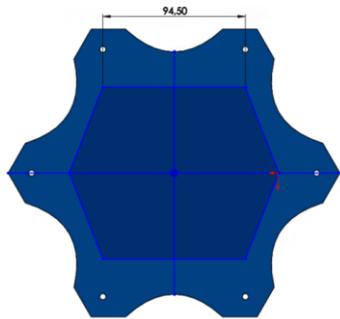


Figure 5. View of the drawn chassis.

A telescopic rod is placed at the center of gravity of the robot chassis. For the interior of the telescopic rod, 5 rollers are designed to engage one another and these rollers are associated with each other using a mate. The task of telescopic rod is to carry the camera. The camera is placed on the other end of this part. For the design of the camera, a solid model with an edge length of 22x22 mm was first created. The formed solid model was given a depth of 22 mm. The designed telescopic rod and camera is shown in Fig. 6 and Fig. 7. respectively.

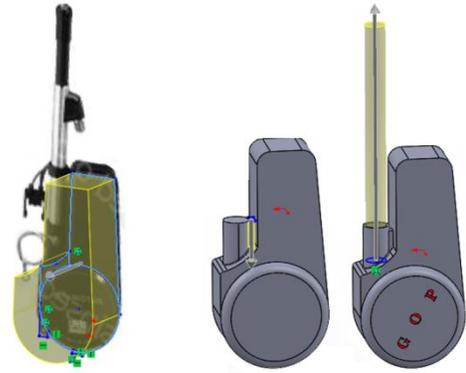


Figure 6. View of the drawn telescopic rod.



Figure 7. View of the drawn camera.

For servo motor rough drawing to provide leg movements, 40 mm height was given and 2 mm "radius" was applied to 4 corners. There are two servo motors on one leg of the robot. The first servo motor is attached directly to the chassis. The second servo motor is attached to the first servo motor by a leg member. The second servo motor has a leg member at the other end. The servo motor and leg design process is shown in Fig. 8 and Fig. 9 respectively.

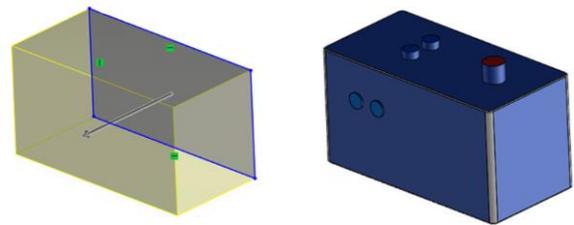


Figure 8. View of the drawn servo motor.

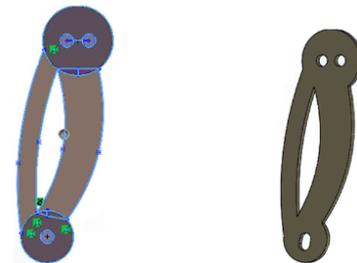


Figure 9. View of the drawn leg.

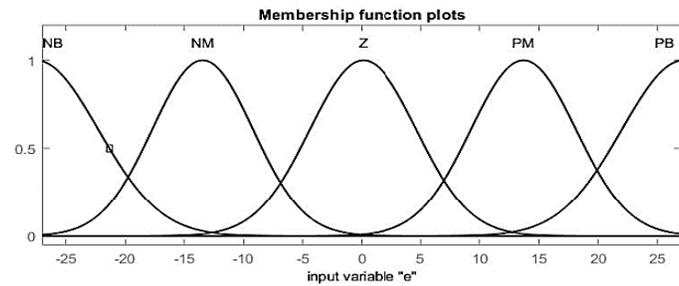
All designed parts of the designed robot are integrated into the assembly tab of the SolidWorks three-dimensional design program. This relationship differs according to the geometry of the contact points of the parts. These mates, which will directly affect the simulation control to be performed on the base of Simulink. Fig.10 shows the assembled robot.



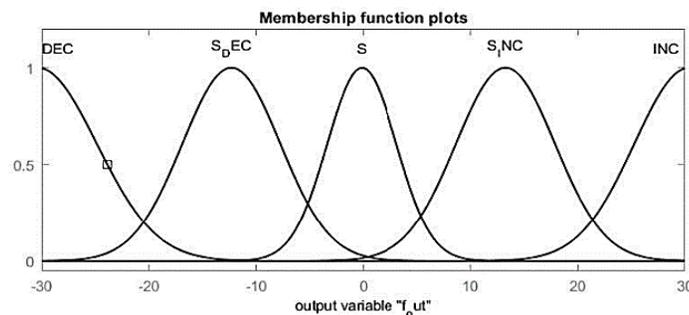
Figure 10. View of the assembled robot.

B. PID and FLC Design

In the FLC design, only the error (e) value was used as the input variable. The membership functions used for e and FLC output (f_{out}) are shown in Fig. 11. As seen, 5 gauss membership functions are used for e and f_{out} .



(a)



(b)

Figure 11. Membership functions used in FLC design.

The rules used for FLC are as follows. NB, NM, Z, PM and PB in these rules and in Fig. 11.a are abbreviations of Negative Large, Negative Medium, Zero, Positive Medium and Positive Large linguistic variables, respectively. DEC, S_DEC, S, S_INC, and INC to the right of the rules and in Fig. 11.b are abbreviations for the linguistic variables Reduce, Slightly Reduce, Wait, Slightly Increase, and Increase, respectively.

1. If (e is NB) then (f_{out} is DEC)
2. If (e is NM) then (f_{out} is S_DEC)
3. If (e is Z) then (f_{out} is S)
4. If (e is PM) then (f_{out} is S_INC)
5. If (e is PB) then (f_{out} is INC)

The same FLC was used for trajectory follow-up controls in the x and y directions. The outputs of both the x -direction and y -direction FLCs are multiplied by 50 gain value.

In the PID control experiments applied in the study, as in FLC experiments, the robot was controlled by two PID controllers on two axes including x and y . Integer order PID is used as PID controller. Adjustment of PID coefficients was done by trial and error method with repeated experiments. The PID coefficients used in the experiments are given in the table below. The outputs of the PID controllers are not multiplied by any gain coefficient (Tab.1).

Table 1. PID controller coefficients on the x and y axes.

Coefficients	PID $_x$	PID $_y$
K_p	10	14
K_i	0.5	0.4
K_d	5	5

IV. RESULTS

Trajectory must be created in order to send the robot to the desired and predetermined points within the working area. The trajectory graph for the robot is given in Fig. 12. According to the graph in Fig. 12, the robot must complete its movement in 150 seconds. It must pass through certain points on the y -axis within 150 seconds. The values in this graph constitute the x or y trajectory reference for the PID and FLC controllers that perform the orbit control of the robot. The motion trajectory of the robot on the x -axis is linear and the reference on the x -axis is zero for 0-150 seconds.

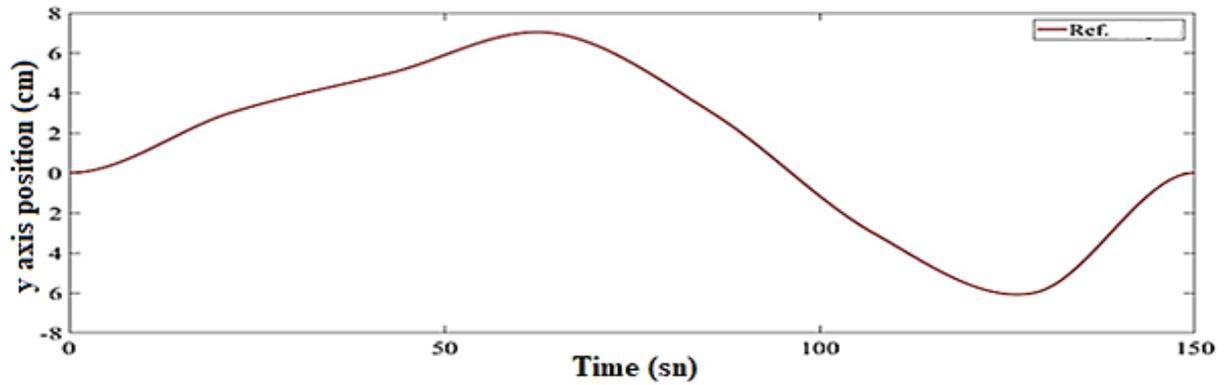


Figure 12. Robot reference trajectory graph.

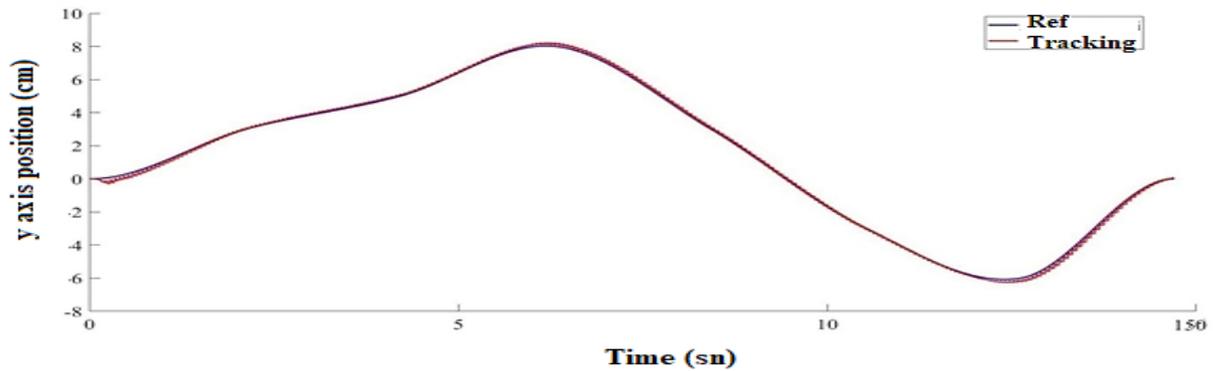


Figure 13. Trajectory graph for the PID controller.

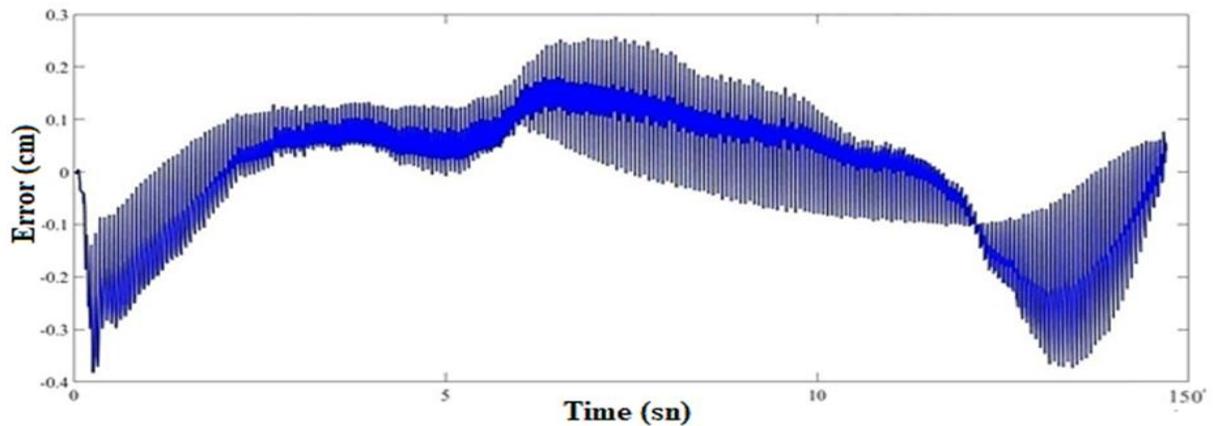


Figure 14. Trajectory tracking error graph on the y-axis for the PID controller.

Simulation experiments were first performed by applying PID controller to the robot and the graphs given below were obtained. The trajectory signal graphs generated on the x and y axes for the PID control method are given in Fig. 13-14.

When the Fig. 13 is examined, it is seen that the trajectory tracking error is higher in the rotation points on the trajectory (upper and lower peaks), but the PID controller can reduce this error in a short time.

For the PID controller, the error values obtained by taking the difference of the y -axis tracking signal generated as a

result of the trajectory signal generated by the robot following the y -axis and the y -axis tracking signal are shown in Fig. 14. When the entire travel time is taken into consideration, the graph shows an error value of 0.3 cm initially generated by the robot in the process of capturing the reference signal. This is due to the friction of the robot feet with the floor and the inertia of the robot itself. Apart from this error, the average trajectory tracking error is seen to be 0.1 on average. This amount of error is a negligible range for a robot moving in terrain and tracking can be said to be a successful follow-up.

In order to better see the trajectory error of the PID controller, a cross section is taken on the trajectory graph on the x axis and this is shown in Fig. 15. Fig. 15 shows that the robot averages the signal towards the end of the said time. The reason for the undulating structure is that the robot is controlled by tripod gaits.

difference of the reference signal and the tracking signal during the stage of controlling the robot on the x axis with PID controller is shown in Fig. 16. In the graph where the movement is shown completely, the error is seen to be within a change range of 0.4 cm. When this graph is analyzed, it is seen that the robot completes the tracking in the x axis with very negligible errors.

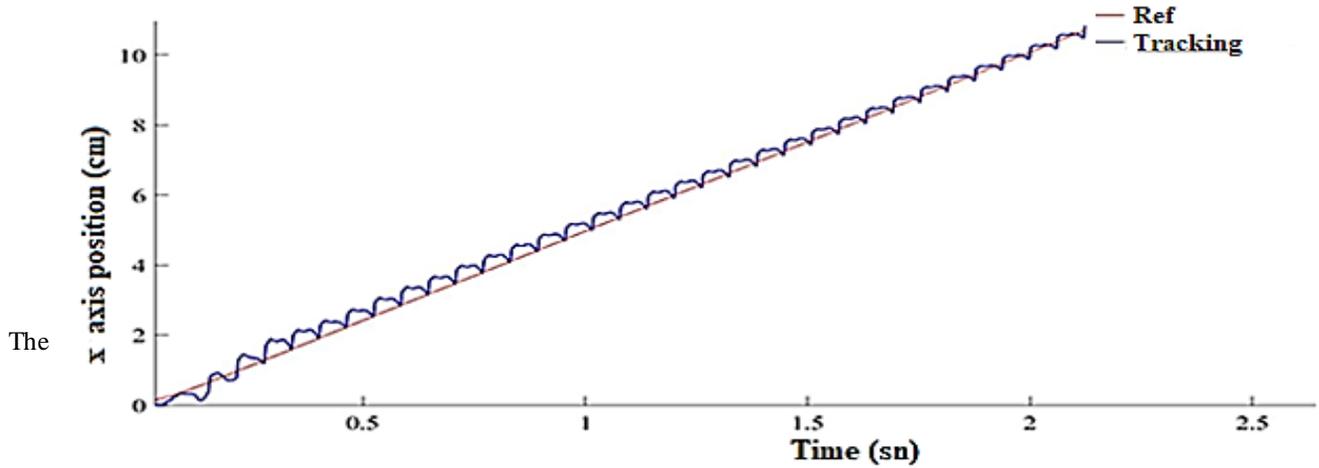


Figure 15. Trajectory graph on the x -axis for the PID controller.

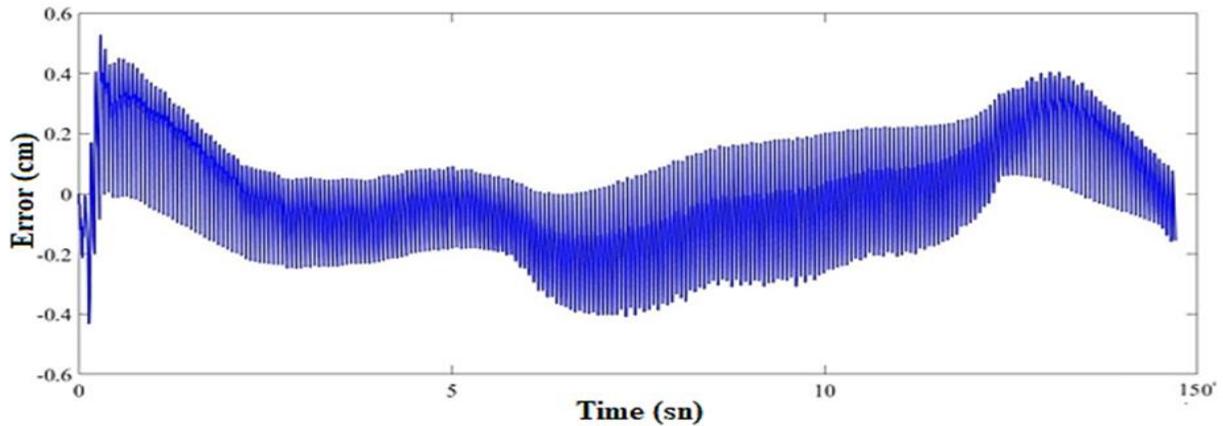


Figure 16. Trajectory tracking error graph on the x -axis for the PID controller.

The zoomed representation of the range of 25-29 sec during control of the robot's x -axis movement for the PID controller is shown in Fig. 17. From this graph, the reference tracking of the robot can be seen more clearly. The reason that the tracking signal is fluctuating is due to the robot's tripod gait motion.

Fig. 18 shows the trajectory tracking performance of FLC. According to Fig. 18, about 10-15 seconds after the start of movement, the robot was able to capture the reference signal and autonomously follow the reference signal. Acceptable trajectory tracking error was obtained along the trajectory.

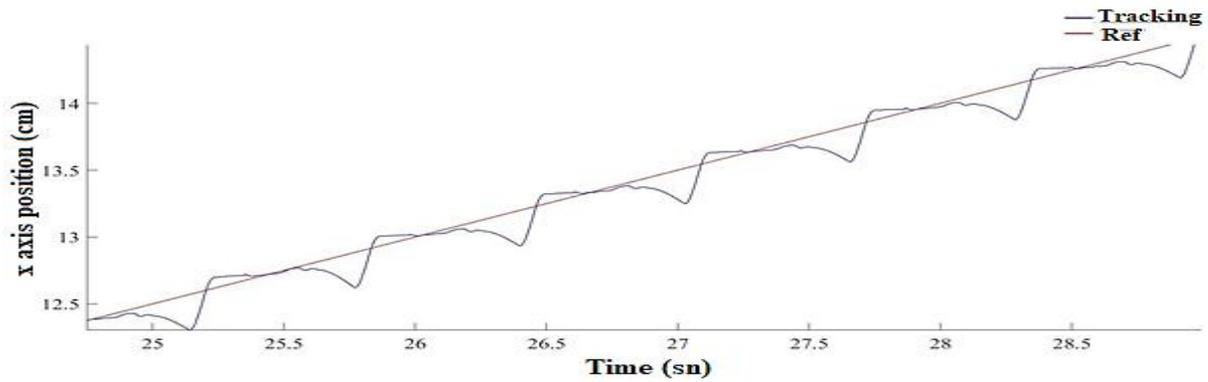


Figure 17. 25-29 sec interval of the reference-follow signal on the x -axis for the PID controller.

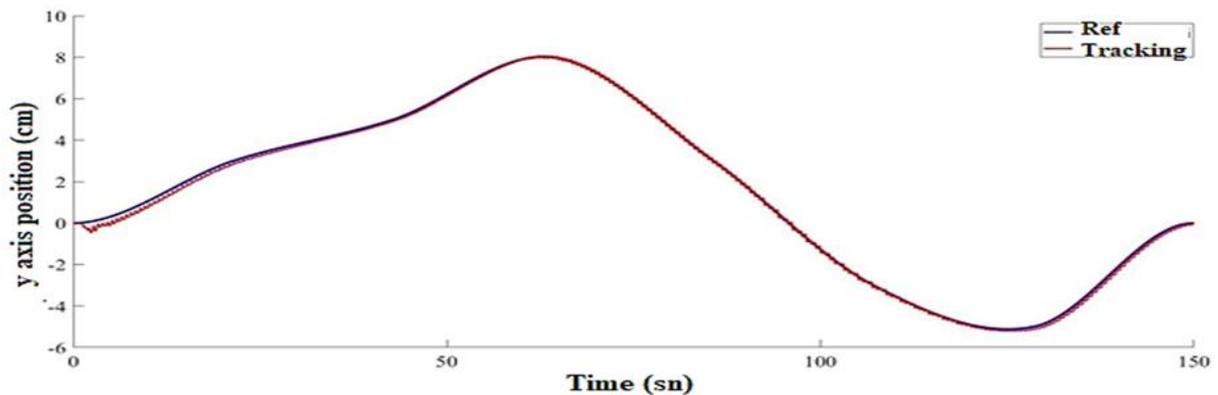


Figure 18. Trajectory graph on the y -axis of the FLC.

Fig. 19 shows the error signal on the y axis made with FLC. It appears that the signal initially oscillates to capture the reference signal and captures the reference value after 10-15 seconds. When the error change interval is examined considering the whole movement time, it is seen that the change range is between 0.1 cm and -0.1 cm. When the results obtained on the y axis of the PID controller are compared, it is seen that the error change interval of the FLC and the error change interval obtained with the PID controller are approximately the same. However, the error change interval at the turning points of the PID controller appears to be slightly larger than the FLC.

Fig. 20 shows the error signal made by the robot in the x -axis reference tracking with FLC. The error signal was obtained by taking the difference between the reference and follow signals. It is seen in the graph that the whole movement time is based on the error range between 0.2 cm and -0.2 cm. When the PID controller is compared with the x -axis error change graph, it is seen that there is more oscillation in PID controller performance than FLC results.

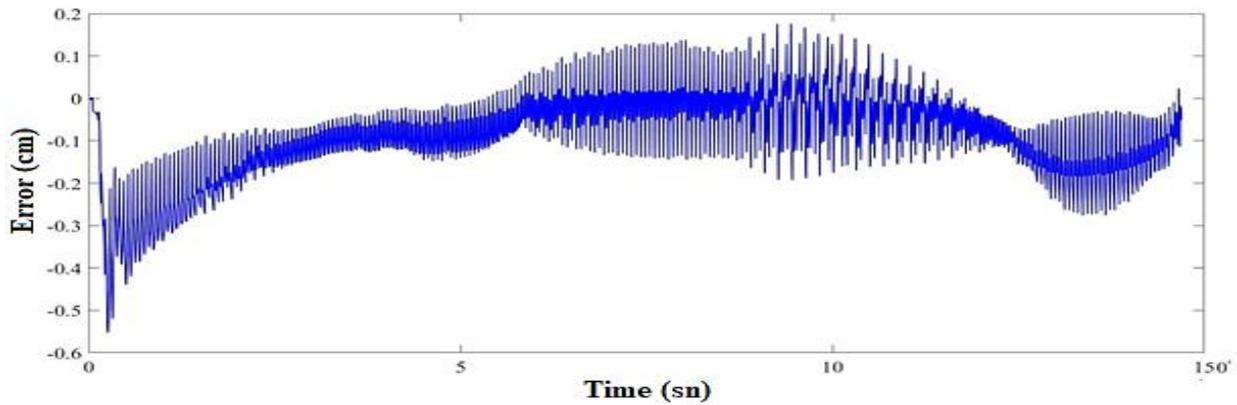


Figure 19. Reference trace error graph on the y-axis of the FLC.

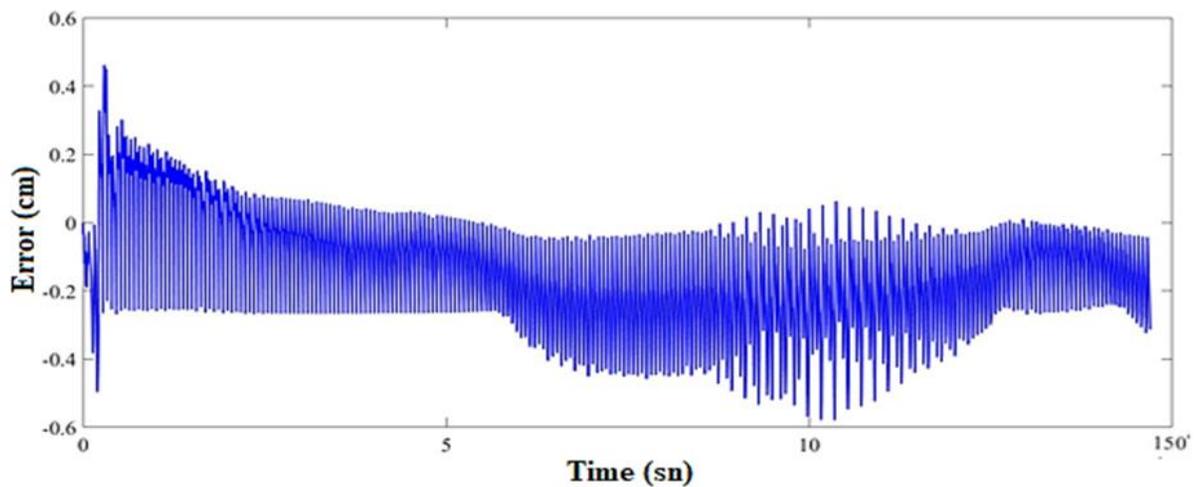


Figure 20. Error graph on the x-axis of the FLC.

V. CONCLUSION

In this study, CAD design of a six-legged spider robot was made in Solidworks program and these CAD designs were transferred to Simulink environment. Robot control simulation studies were performed in Simulink environment. The robot is intended to move autonomously on a particular trajectory. Robot motion trajectory is obtained by curve fitting methods. PID and FLC controllers were used for robot trajectory controls and robot trajectory motion performances of the two controllers were compared. In the study, during the controls, linear reference signal on x axis and an orbit on y axis were used in two types of controllers. Follow-up graphs were obtained during robot movement. In the analysis studies, the total weight of the robot with 4 kg was considered.

As a result of the analysis studies, it was observed that this robot, which was designed and analyzed, could follow a trajectory defined by a mass of 4 kg with an error value of approximately 1 mm. When FLC and PID performances are examined, it is concluded that the trajectory is followed with approximately the same error rates. During the trajectory tracking, the FLC controller performs better than the PID

controller at the turning points of the reference trajectory curve of the robot. It would be useful to convert the design outputs to a real physical hardware and edit the motion control algorithms on a microcontroller system to create a physical prototype of the designed robot and to test the motion control methods under real environment conditions.

REFERENCES

- [1] G. Gürgüze, İ. Türkoğlu. "Kullanım Alanlarına Göre Robot Sistemlerinin Sınıflandırılması." *Fırat Üniversitesi Müh. Bil. Dergisi* 31(1), pp. 53-66, 2019.
- [2] F. Tedeschi, G. Carbone. "Design Issues for Hexapod Walking Robots." *Robotics* 2014, 3, pp. 181-206; doi:10.3390/robotics3020181
- [3] D. Chávez-Clemente. "Gait Optimization for Multi-legged Walking Robots, with Application to a Lunar Hexapod." Ph.D. Thesis, Stanford University, California, CA, USA, 2011.
- [4] G. Carbone, M. Ceccarelli. "Legged Robotic Systems. Cutting Edge Robotics" ARS Scientific Book, Vienna: pp. 553-576, 2005.
- [5] S. E. Hamamci. "A New PID Tuning Method Based on Transient Response Control." *Balkan Journal of Electrical & Computer Engineering*, Vol.2, No.3, 2014 pp 132-138
- [6] Caponetto, R., Fortuna, L., Porto, D. "Parameter Tuning of A Non-Integer Order PID Controller." 15th International Symposium on Mathematical Theory of Networks and Systems, Notre Dame, Indiana, 2002.

- [7] A. Ates, B. B. Alagoz, G.T. Alisoy, C. Yeroglu and H. Z. Alisoy, "Fuzzy Velocity and Fuzzy Acceleration in Fractional Order Motion." *Balkan Journal of Electrical & Computer Engineering*, Vol 3, No 2, 2015, pp 98-102, DOI:10.17694/bajece.52354
- [8] A. Sarabakha, C. Fu, E. Kayacan. "Intuitbeforetuning: Type-1 and type-2 fuzzy logic controllers." *Applied Soft Computing*, Vol 81, 2019. <https://doi.org/10.1016/j.asoc.2019.105495>
- [9] O. Castillo, L. Amador-Angulo, J.R. Castro, M. Garcia-Valdez, A comparativestudy of type-1 fuzzy logic systems, interval type-2 fuzzy logic systemsand generalized type-2 fuzzy logic systems in control problems, *Inform.Sci.* 354 (2016) 257–274, <http://dx.doi.org/10.1016/j.ins.2016.03.026>.
- [10] L. Cervantes, O. Castillo. "Type-2 fuzzy logic aggregation of multiple fuzzycontrollers for airplane flight control." *Inform. Sci.* 324 (2015) 247–256, <http://dx.doi.org/10.1016/j.ins.2015.06.047>.
- [11] R.-E. Precup, H. Hellendoorn, "A survey on industrial applications of fuzzycontrol." *Comput. Ind.* 62 (3) (2011) 213–226, <http://dx.doi.org/10.1016/j.compind.2010.10.001>.
- [12] A. Celikyilmaz, I.B. Turksen. "Modeling Uncertainty with Fuzzy Logic: With Recent Theory and Applications." first ed., Springer-Verlag Berlin Heidelberg, 2009, <http://dx.doi.org/10.1007/978-3-540-89924-2>.
- [13] T. Kumbasar, H. Hagrass. "Big Bang-Big Crunch optimization based interval type-2 fuzzy PID cascade controller design strategy." *Inform. Sci.* 282 (2014) 277–295, <http://dx.doi.org/10.1016/j.ins.2014.06.005>.
- [14] C. Fu, A. Sarabakha, E. Kayacan, C. Wagner, R. John, J.M. Garibaldi. "Input uncertainty sensitivity enhanced nonsingleton fuzzy logic controllers forlong-term navigation of quadrotor UAVs." *IEEE/ASME Trans. Mechatronic*

the University of Inonu, in 2017. From 2007 to 2018, he was a Lecturer in Zile Vocational School. Since 2018, he has been an Assistant Professor with the Electrical-Electronic Engineering Department, Tokat Gaziosmanpasa University. He is the author of 9 articles, 9 inventions. His research interests include automatic control applications, robotic, fuzzy logic, artificial neural networks.

BIOGRAPHIES



S. AYDIN received the B.S. degrees in mechatronic engineering from the University of Tokat Gaziosmanpasa University, in 2017. He currently studying M.S. degree in Tokat Gaziosmanpasa University, Graduate School of Natural and Applied Sciences from 2018. His research interests include robotic, CAD design of mechanical systems.



LEVENT GÖKREM was born in Tokat, TURKEY in 1970. He received the B.S. degree in Department of Electronics and Computer Education from Gazi University, Turkey, in 1993. He received the M.S. degree in Electronics and Computer Education from the Sakarya University Turkey, in 2003 and the Ph.D. degree in Industrial Technology Education from Gazi University, Ankara, in 2009. He has been an Assistant Professor with the Department of Mechatronics Engineering, Tokat Gaziosmanpasa University, Turkey. His research areas are Internet of Things, Image processing, robotic and wireless communication.



M. S. CAN received the B.S. degree in electrical-electronic engineering from the University of Nigde, in 2000 and M.S. degrees in electrical-electronic engineering from the University of Sutcuimam, in 2010 and the Ph.D. degree in electrical-electronic engineering from

Phase Changing Material Usage to Increase the Efficiency of Photovoltaic Panels

H. KARAKAYA, İ.E. ŞEN

Abstract— The energy demand from renewable energy sources has been proportionally increased due to population growth. In renewable energy production, not only energy production but also efficient use is crucial. In recent years, photovoltaic panels, which are increasingly used and highly important in terms of renewable energy production, have warmed on the surface due to the radiation energy emitted by the sun, and this warmth reduces the efficiency of the panel. In this study, it is aimed to reduce this efficiency loss by using phase changing material with latent heat which is one of the thermal energy storage methods. One of the two solar panels placed parallel to the ground was filled with phase changing material. With this, it is aimed to cool photovoltaic panel by using calcium chloride hexahydrate phase changing material; and then radiation, air temperature, current, voltage, power, surface temperature, efficiency graphs were obtained. As a result of simultaneous measurements, the increase in yield was found to be 2.95%.

Index Terms— Efficiency, Latent heat, Phase change material, Photovoltaic panel.

I. INTRODUCTION

INTENSITY OF solar radiation, surface cleaning & temperature, and angle of inclination of the photovoltaic (PV) cell are the specific physical parameters that affect the efficiency of PV panel.

Electricity is produced by PV panels which only convert 5-25% of solar irradiance into energy [1], and the remaining radiation is transformed into heat energy [2] which increases PV panel temperature [3]. Therefore, PV panel efficiency decreases, while the surface temperature increases [4] due to the internal charge carrying combination rates [5], which is the cause for reduction in maximum power production. As a consequence of this, a *suitable* cooling method that consists of

HAKAN KARAKAYA, is with Department of Mechanical Engineering University of Batman University, Batman, Turkey, (e-mail: hakan.karakaya@batman.edu.tr).

 <https://orcid.org/0000-0001-9242-6233>

IZZETTIN ENES SEN, is with Department of Mechanical Engineering University of Batman University, Batman, Turkey, (e-mail: ienessen@gmail.com).

 <https://orcid.org/0000-0002-5791-217X>

Manuscript received November 29, 2019; accepted January 27, 2020.
DOI: [10.17694/bajece.666891](https://doi.org/10.17694/bajece.666891)

active (directly reducing the temperature of PV panel) and passive method (methodology that works as a heat sink and takes the extra heat from the panel) should be used in order to increase the electrical efficiency of PV panels. Passive cooling method for PV panels is a widely used technique to sustain the temperature of PV module at a level which attains higher efficiency since active cooling method has rather more maintenance and operating problems. On the other hand, the Phase Change Material method is a passive temperature controlling technique of PV modules which does not need high operational cost; on the contrary, it uses less energy density [6].

Recent studies have revealed that, researchers have endeavored to build up a thermal model to evaluate the system of PV operating temperature considering fine details of heat transfer. In their study, Ceylan et al. (2014) experimentally investigated what kind of an effect cooling can have on the system [7]. When they carried out cooling operation by installing a spiral heat changer at the back of a PV panel, they found that the efficiency was 13%; however, when they did not perform cooling operation, it was 10%. In addition to this, they observed that as the sun radiation increased, temperature values decreased.

D'Avignon and Kummert (2016) conducted an experimental study in order to evaluate the performance of phase changing substance storage tank under various operating conditions on a dynamic test bench. There were plates like FDM capsule piles where the heat transfer liquid could move through the investigated horizontal storage tank. They observed significant differences in behaviors of FDMs under the same test conditions (Extreme cooling until different temperatures and different phase changing temperature) [8].

Kabeel et al. (2016) performed their study in the city of Tanta, Egypt using paraffin wax, a phase changing substance. They carried out the investigation experimentally employing air heaters with flat and V corrugated plates [9]. They studied the parameters that had influence on thermal performance of solar air heaters with and without substances. They concluded that the daily efficiency of V corrugated solar heater using FDM was 12% higher than that of the FDM not using solar heater, and that the daily efficiency was 15% higher when the flat plate was used.

Konuklu et al. (2017) produced diatomite phase changing material composites using diatomite raw material. As a result of differential scanning of the prepared composites, melting and freezing temperature of the synthesized composites were

found as 41,97°C and 45,77°C by using calorimeter, respectively, while latent heat storage capacity was found as +45 and -44 J/g. They concluded that the above mentioned substances could be used for thermal energy storage [10].

In their study, Stropnik and Stritih (2016) conducted research on how to increase the efficiency of PV cells by using phase changing material. As a result of this study, they concluded that in Ljubljana, the photovoltaic cell, in which phase changing material was used, produced 7.3% more energy in a year [11].

In the study they carried out, Sarhaddi et al. (2017) investigated the heat storage performance of 2 weir type cascade connected solar power systems with phase-changing substances on sunny and semi-sunny days and without phase-changing substances; and they performed energy and exergy analyses. They found 74.35% energy and 8.59% exergy efficiency of solar energy with phase change substance for a half sunny day. They concluded that the highest irreversibility was obtained with the absorber plate [12].

Kant et al. (2016) predicted that the surface temperature of the panel may decrease due to the high latent heat storage capacity of the PV panels used with phase change materials. They found that high wind velocity and tilt angle play an important role in decreasing the temperature of the panel as a result of the heat transfer modeling with the fluid dynamics program they employed [13].

Guarino et al. (2017) concluded that energy performance was appropriate as a result of empirical studies, and they supported them with simulation by embedding phase-changing substance in the southern façade of a building in cold climatic conditions [14].

Peng et al. (2017) investigated the relationship between surface temperature and efficiency of PV panels. They found that the efficiency of PV panels could reach up to the levels of 47% and the payback time might decrease by 20% as a result of experimental studies [15].

Andre et al. (2018) suggested that photovoltaic aqueous cooling system may have a positive effect on heat transfer. They observed that the efficiency of water-cooled panels increased by around 12% in the study they carried out [16].

Chen et al. (2019) investigated the use of ventilation systems with phase changing material in thermal energy storage in their study. In the study, an office room in Beijing vicinity was chosen for research. Thermal model, heat transfer model and performance curve of the room were found out. When they provided cooling by using a phase-changing substance, the achievement from electrical energy was between 16.9% and 50.8%, but when cooling was not provided, it was between 9.2% and 33.6% [17].

Fayaz et al. (2019) made numerical analyses and compared them experimentally with the help of three dimension COMSOL software in their study to improve heat transfer and performance in photovoltaic thermal systems. They provided cooling with phase changing material with an input at 27 °C temperature with different flow rates and achieved maximum electrical efficiency 12.4% as numerical and 12.28% as experimental. Similarly, in the case when they carried out the

analyses on the PV panel with phase changing substance, they achieved 12.75% and 12.59% electrical efficiency in experimental and numerical cases, respectively [18].

In our study, PV panel efficiency was provided using phase changing substance thus optimizing the surface temperature of PV panel to decrease efficiency loss because of high temperatures on PV panel surface [19].

II. GUIDELINES FOR GRAPHICS PREPARATION AND SUBMISSION

Experiment set was placed on Batman University central campus terrace during August 2019. The irradiation values varied between 11-919 W/m². One multimeter, 2 solar panels, 9 kg phase changing material, two recorders, 8 K type thermocouples, 1 laptop and 6 resistors were used in the experiment. Two identical 125 W solar panels with 114 cm length and 67 cm width with an open circuit voltage of 22.10 V and a short circuit current of 7.30 A were used in the experiments. Calcium chloride hexahydrate imported from abroad with a melting temperature of 30°C and a melting temperature of 140 kJ / kg and a density of 1.71 g / ml was used in the solar panels [8]. Elimko 680 temperature data logger was used for measurements. GRAPHTEC midi LOGGER GL240 Series was used as voltage recorder. Measurements were carried out with ± 0.1% error.

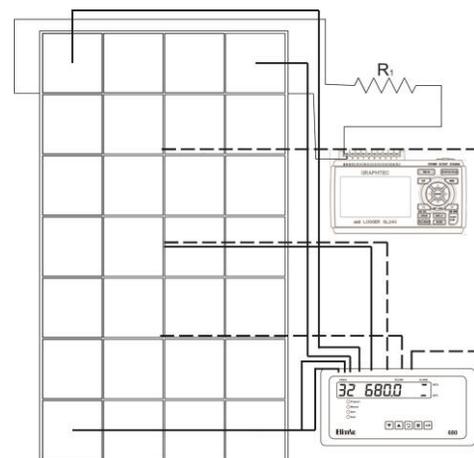


Fig.1. Experiment set

Temperature values were obtained from 5 different points indicated by bold lines in the Fig. 1. The temperature values of the phase changing substance were obtained from 3 different points, which are indicated by thin lines in the Fig.1. As can be seen in Fig. 1, a resistor of 3.3 Ω, shown as R₁, was used to obtain maximum power in photovoltaic panel.

III. MATHEMATICAL BACKGROUND

The latent heat is calculated as follows:

$$Q = mL \quad (1)$$

Here: Q – is the latent heat of the phase changing substance;

m – is the mass of the phase changing substance;
 L – is the specific latent heat coefficient of the material;

The voltage is calculated as follows:

$$V = IR \tag{2}$$

Here: V – is the voltage of the photovoltaic panel;

I – is the current of the photovoltaic panel;

R – is the resistance;

The power is calculated as follows:

$$P = IV \tag{3}$$

Here: P – is the power of the photovoltaic panel;

I – is the current of the photovoltaic panel;

V – is the voltage of the photovoltaic panel;

The efficiency is calculated as follows:

$$\eta = \frac{P_{out}}{P_{in}} * 100$$

Here: η – is the efficiency of the photovoltaic panel;

P_{out} – is the output power of the photovoltaic panel which comes from solar irradiation;

P_{in} – is the input power of the photovoltaic panel which is created in itself; (4)

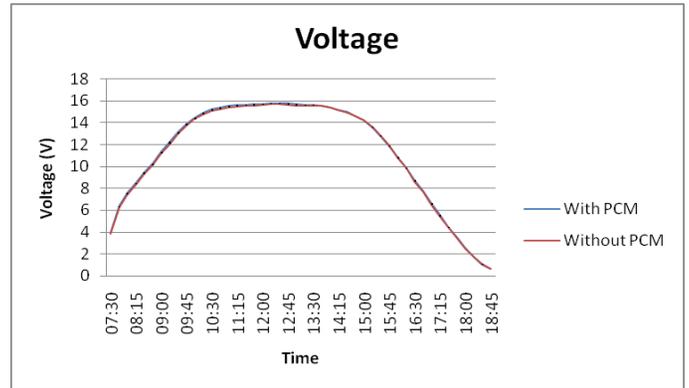


Fig.5. Voltage graph

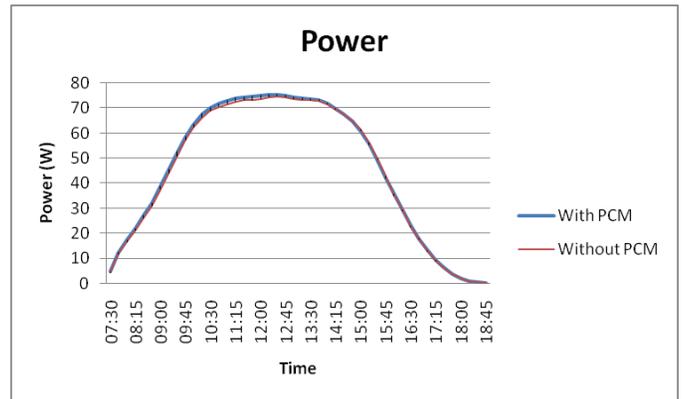


Fig.6. Power graph

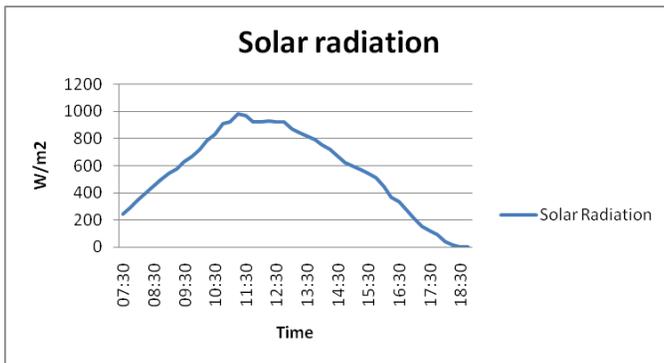


Fig.3. Air temperature graph

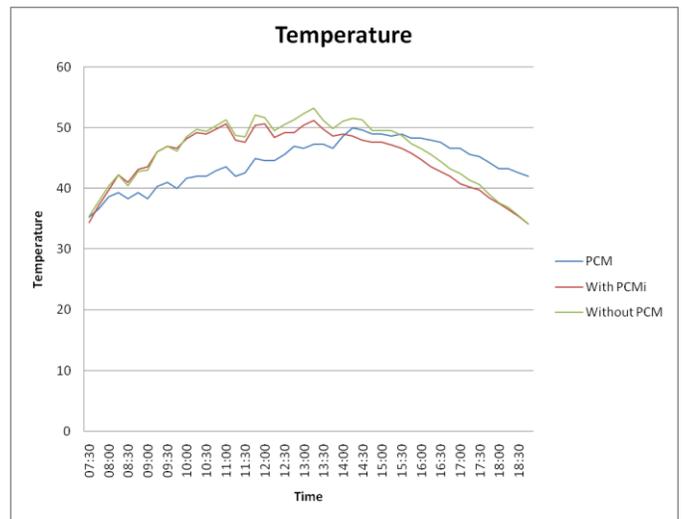


Fig.7. Temperature graph

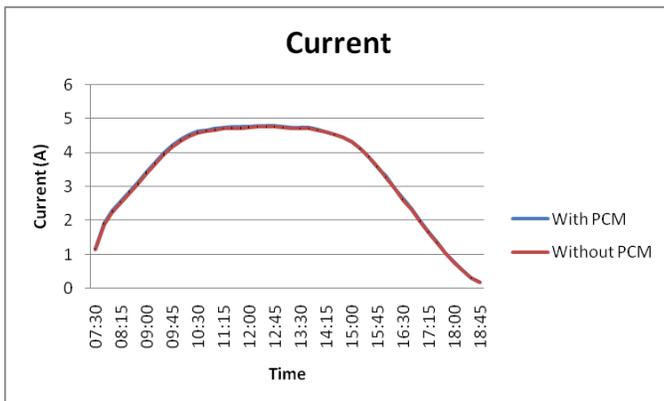


Fig.4. Current graph

IV. DISCUSSIONS

The graph of the radiation values is shown in Fig.2. The highest radiation value was measured as 981 W/m^2 at 12:30. As with the power, in voltage and current graph, the maximum values were measured at 12:30, which demonstrates that this graph supports the other graphs. As shown in Fig.3, the maximum air temperature is 32.47°C at around 13:30 and the minimum air temperature is 24.62°C at 7:30. In Fig. 4, as the sun rises, the sun rays fall more perpendicular on to the panel surface, the production capability of the panel increases and it can produce higher current values. The maximum current value obtained using phase changing material is 4.78 A at 12:30. However, 4.75 A was obtained at the same time without phase change material. As shown in the Fig.5, the voltage rises were observed as a result of the solar rays falling perpendicular to the PV panel surface. The maximum voltage value obtained using the phase changing material is 15.30 V at 12:30. However, 15.7 V is obtained at the same time without phase change material. The power graph that is multiplied by current; and voltage graphs increase with increasing of solar radiation on to the surface of the PV panel. As can be seen in the Fig. 6, the power reaches the maximum value at noon; and then starts to decrease again. The maximum power obtained using the phase changing substance is 75.45 W at 12:30. 74.69 W is obtained at the same time without phase change substance. The graph in Fig. 7 demonstrates the time-dependent temperature. The red and green lines indicate the surface temperature of the PV panel. At 13:15, the maximum temperature of the PV panel surface with phase changing substance is 51.2°C , whereas the other PV panel without phase changing substance is 53.26°C . The temperature of the phase-changing substance is measured at a maximum of 50°C at 14:15. 9 kg phase changing substance used in the experiment melted from 8:00 to 9:00.

V. CONCLUSION

Nowadays, fossil fuels are insufficient to provide energy for the increasing energy demand. As a solution to this issue, energy production, which consists of renewable energy sources, known as eco-friendly, is increasing rapidly. Not only energy production provides a solution for the growing energy demand as a consequence of population growth, but also using the same energy source with more efficient way is necessary. PV panels, which use clean renewable solar energy, and therefore eco-friendly, are efficient in electricity generation. For these reasons, in this study, it is aimed to decrease the high temperature formed on the PV panel surface by using phase changing substance to provide higher efficiency. As a result of the experimental studies, the PV panel was cooled by using phase changing material called calcium chloride hexahydrate; and radiation, air temperature, current, voltage, power, surface temperature, yield graphs were obtained. The efficiency increase in PV panel was measured as 2.95% by using phase changing substance. The maximum increase for current and voltage was equal to 1.46% . The result can be a small achievement in this experiment when we consider small scale increase in efficiency of PV panels. However, it will play an important role in large-scale investments.

ACKNOWLEDGMENT

This study was supported by the scientific research Coordinatorship of BATMAN University with the project BTÜ BAP 2019-YL-03.

References

- [1] K. Kant, A. Shukla., A. Sharma, P.H. Biwole. "Thermal response of polycrystalline Silicon photovoltaic panels: numerical simulation and experimental study", *Sol. Energy*, vol. 134, 2016, pp. 147–155.
- [2] P. Atkin, M.M. Farid. "Improving the efficiency of photovoltaic cells using PCM infused graphite and aluminum fins", *Sol. Energy*, vol. 114, 2015, pp. 217–228.
- [3] C.S. Malvi, D.W. Dixon-Hardy, R. Crook. "Energy balance model of combined photovoltaic solar-thermal system incorporating phase change material", *Sol. Energy*, vol. 85, 2011, pp. 1440–1446.
- [4] C.S. Malvi, D.W. Dixon-Hardy, R. Crook. "Energy balance model of combined photovoltaic solar-thermal system incorporating phase change material", *Sol. Energy*, vol. 85, 2011, pp. 1440–1446.
- [5] B. Zhao, W. Chen, J. Hu, Z. Qiu, Y. Qu, B. Ge "A thermal model for amorphous silicon photovoltaic integrated in ETFE cushion roofs", *Energy Convers. Manage.*, vol. 100, 2015, pp. 440–448.
- [6] A. Hasan, S.J.J. McCormack, M.J.J. Huang, B. Norton. "Evaluation of phase change materials for thermal regulation enhancement of building integrated photovoltaics", *Sol. Energy*, vol. 84, 2010, pp. 1601–1612.
- [7] I. Ceylan, A. E. Gürel, H. Demircan, & B. Aksu, "Cooling of a photovoltaic module with temperature controlled solar collector" *Energy and Buildings*, 72, 96–101, 2014.
- [8] K. D'Avignon, M. Kummert. "Experimental Assessment of a Phase Change Material Storage Tank", *Applied Thermal Engineering*, vol. 99, 2016, pp. 880–891.
- [9] A. E. Kabeel, A. Khalil, S. M. Shalaby, & M. E. Zayed "Experimental investigation of thermal performance of flat and v-corrugated plate solar air heaters with and without PCM as thermal energy storage" *Energy Conversion and management*, 113, 264–272, 2016.
- [10] Y. Konuklu, O. Ersoy, H. Paksoy, S. Evcimen, S. Celik, O. Toraman. "Production of Diatomite/Phase Changing Material Composites As Thermal Energy Storage Material. Omer Halisdemir University Journal of Engineering Sciences", vol. 6, no: 1, 2017, pp. 238-243.
- [11] R. Stropnik and U. Stritih. "Increasing the efficiency of PV panel with the use of PCM", *Renewable Energy*, vol. 97, issue C, 2016, pp. 671–679.
- [12] F. Sarhaddia, F.F. Tabrizi, H.A. Zoori, S.A. Hossein, S. Mousavi. "Comparative study of two weir type cascade solar stills with and without PCM storage using energy and exergy analysis", *Energy Conversion and Management*, vol. 133, 2017, pp. 97–109.
- [13] K. Kant., A. Shukla, A. Sharma, P.H. Biwole. "Heat transfer studies of photovoltaic panel coupled with phase change material", *Solar Energy*, vol. 140, 2016, pp. 151–161.
- [14] F. Guarino, A. Athienitis, M. Cellura, D. Bastien. "PCM Thermal Storage Design in Buildings: Experimental Studies and Applications to Solaria in Cold Climates", *Applied Energy*, vol. 185, 2017, pp. 95–106.
- [15] Z. Peng, M.R. Herfatmanesh, Y. Lu, "Cooled solar PV panels for output energy efficiency optimisation", *Energy Conversion and Management*, vol. 150, 2017, pp. 949–955.
- [16] F. Andrea, J.F.P. Fernandes, P.J.C. Branco, "Demonstration Project of a cooling system for existing PV power plants in Portugal", *Applied Energy*, vol. 211, 2018, pp. 1297–1307.
- [17] X. Chen, Q. Zhang, Z. J. Zhai, X. Ma. "Potential of ventilation systems with thermal energy storage using PCMs applied to air conditioned buildings", *Renewable Energy*, vol. 138, 2019, pp.39–53.
- [18] H. Fayaz N.A. Rahim, M. Hasanuzzaman, R. Nasrin, A. Rivai. "Numerical and experimental investigation of the effect of operating conditions on performance of PVT and PVT-PCM," *Renew. Energy*, vol. 143, 2019, pp. 827–841.
- [19] I.E. Sen, "Experimental Investigation of The Effect Of Phase Changing Substance Use In Photovoltaic Panels On Efficiency" Batman University Graduate School of Natural and Applied Science, Unpublished MasterThesis, 2019.

BIOGRAPHIES



HAKAN KARAKAYA was born on 20.12.1979 in Çelikhan district of Adıyaman/Turkey. After successfully completing Elazığ Gazi Anatolian Technical High School Mechanical Department in 1997, he enrolled in Fırat University Technical Education Faculty Mechanical Education Department in 1998 for his undergraduate education.

After completing his undergraduate education in 2002, he started his graduate studies at Fırat University, Institute of Science and Technology, Mechanical Education/Energy. After completing his master's degree in 2005, he started his PhD studies in the same field. After completing his doctorate in 2010, he started working as an Assistant Professor in Batman University, Department of Mechanical Engineering. He is also working as the manager in Vocational school of Beşiri organized industrial zone. Hakan Karakaya is married and has one child.



İZZETTİN ENES SEN was born in 1991. After successfully completing Bahçeşehir University Energy Systems Engineering department in 2014, he enrolled in Batman University Mechanical Engineering Energy Department and graduated in 2019.

Classification of EEG Signals in Depressed Patients

S. G. ERALDEMİR, Ü. KILIÇ, M. KELEŞ, M. E. DEMİRKOL, E. YILDIRIM, L. TAMAM

Abstract—Electroencephalography (EEG) are electrical signals that occur in every activity of the brain. Investigation of normal and abnormal changes that take place in the human brain using EEG signals is a widely used method in recent years. The World Health Organization (WHO) states that one of the most important health problems in today's society is depressive disorders. Nowadays, various scales are used in the diagnosis of depressive disorder in individuals. These scales are based on the declaration of the individual. In recent studies, EEG has been used as a biomarker for the diagnosis of depression. In this study, EEG signals from 30 patients with clinical depressive disorder have been recorded. EEG signals have been collected for 1 minute with eyes open and closed. The collected data have been divided into attributes by continuous wavelet transform which is used in many studies in processing non-stationary signals such as EEG. Obtained attributes have been classified with kNN classification method. As a result, it was observed that EEG signals, collected from subjects with depression while eyes are open and closed, can be classified with an accuracy of 91.30%.

Index Terms—EEG, Depressive Disorders, kNN, Wavelet Transform

SERVER GÖKSEL ERALDEMİR, is with Department of Iskenderun Vocational School of Iskenderun Technical University, HATAY, Turkey, (sgoksel.eraldemir@iste.edu.tr).

 <https://orcid.org/0000-0003-0835-2601>

ÜMIT KILIÇ, is with Department Faculty Engineering of Adana Alparslan Turkes of Science and Technology University, Adana,Turkey, (ukilic@atu.edu.tr)

 <https://orcid.org/0000-0001-8067-6024>

MÜMİNE KAYA KELEŞ, is with Department Faculty Engineering of Adana Alparslan Turkes of Science and Technology University, Adana,Turkey, (e-mail: mkaya@atu.edu.tr).

 <https://orcid.org/0000-0001-8414-1713>

MEHMET EMİN DEMİRKOL, is with Department Faculty Medicine of Çukurova University Faculty, Adana,Turkey, (e-mail: emindemirkol@gmail.com).

 <https://orcid.org/0000-0003-3965-7360>

ESEN YILDIRIM, is with Department of Electrical-Electronics Engineering, Adana Alparslan Turkes Science and Technology University, Adana, Turkey, (e-mail: eyildirim@atu.edu.tr)

 <https://orcid.org/0000-0002-4650-1164>

LUT TAMAM, is with Department Faculty Medicine of Çukurova University Faculty, Adana,Turkey, (e-mail: ltamam@gmail.com).

 <https://orcid.org/0000-0002-9750-7531>

Manuscript received October 11, 2019; accepted December 23, 2019.

DOI: [10.17694/bajece.631951](https://doi.org/10.17694/bajece.631951)

I. INTRODUCTION

IN ANY movement that occurs in the human body, electrical activity occurs in nerve cells in the brain. These activities are recorded with the help of sensors, and the recorded electrical data are called EEG. The frequency range of the EEG data is between 0.5-100 Hz and the amplitude is between 1 and 100 μ V. EEG data are recorded with electrodes placed on the scalp according to the international 10-20 system. Depression was considered by WHO to be the second most important health problem in 2020 that threatened people's lives [1]. Today, various scales are used in the diagnosis of depressive disorder in individuals. The application of various diagnostic scales used in the diagnosis of depression is based on the individual's declaration. In addition to the individual declaration, in recent studies, EEG signals have been used as a biomarker for the diagnosis of depression [2,3]. Also, EEG data are used in the diagnosis of depressive disorders as well as in investigating the effects of drugs used. In 2014, in the study of Erguzel et al. [1], EEG signs were used to determine the amount, level and effect of the drugs used in the treatment of major depressive disorders [1]. In the study, patients with depressive disorder and drug therapy were divided into two groups as responders and non-responders and EEG signals were collected. As a result of this procedure, the EEG data of the responding group and the non-responding group were classified with an accuracy rate of 85.45%. Mallikarjun et al. [4] studied the classification of EEG signals, which are considered to be one of the symptoms of depressive disorder and depressive symptoms of the patients with depressive disorder. In their study, EEG data of 240 patients with sleep problems and 60 patients who needed alcohol were classified and 88.32% accuracy rate was obtained. In the studies conducted recently, EEG signals of individuals diagnosed with Major Depressive Disorder are examined, and it is tried to obtain meaningful results from these signals [5,6]. In this study, EEG signals were collected on a voluntary basis from the patients whose depressive disorder were diagnosed by using the scales applied in the Psychiatry Department of Çukurova University Medical Faculty clinically. These signals are collected by means of electrodes placed in accordance with the international 10/20 system. The aim of this study was to determine whether the EEG data collected from the patients with depressive disorder in the eyes open and closed state could be differentiated from each other.

II. METHODS

In the scope of this study, 30 patients diagnosed as having depressive disorder who were treated in Balcalı Hospital Psychiatry Department of Çukurova University Medicine Faculty were included in the study. In addition to EEG measurements, sociodemographic data form prepared by the researchers, Psychological Pain Scale and Beck Depression Scale were fulfilled by the participants.

A. Psychological Pain Scale

It is a five-point Likert-type self-report scale developed by Holden et al. Nine items of the scale assess the frequency of psychological pain and four items evaluate its intensity. Responses to the scale are in the range between never-always or strongly disagree-strongly agree. It was observed that the scale successfully differentiated suicide attempters and non-attempters [7]. A high score from the scale indicates that

psychological pain is greater. In the Turkish validity and reliability study, Cronbach's alpha value was 0.98 [8].

B. Beck Depression Scale

It is a quartet Likert type self-report scale consisting of 21 items developed by Beck. The higher the score, the greater the severity of depression. In the Turkish validity and reliability study, Cronbach's alpha value was 0.80 [9].

C. Collection of EEG Data

EEG data from 30 patients with clinical depressive disorder were recorded with eyes open and closed. During the recording process, the patient was asked to keep his eyes closed and open for 1 minute according to the voice command. Figure 1 is a representation of recording process.

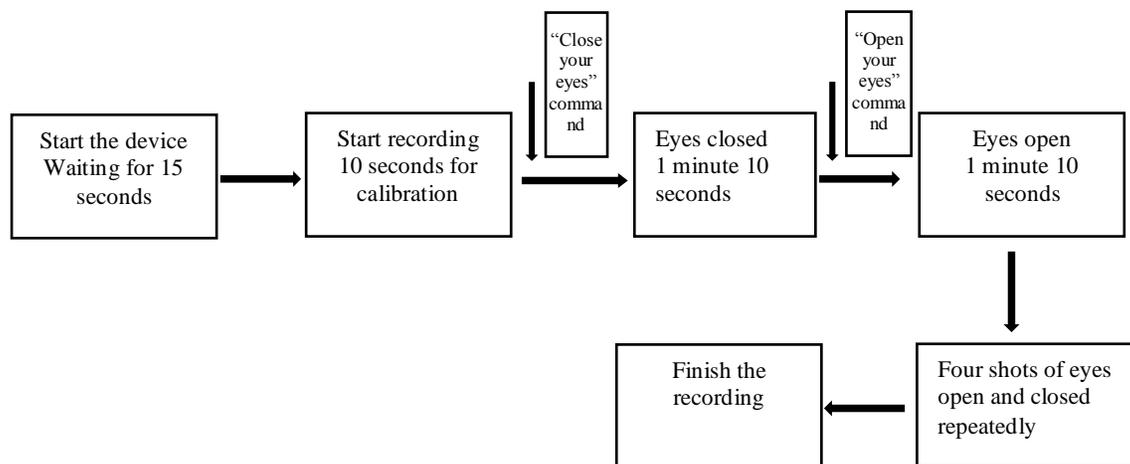


Figure 1. Representation of recording process.

Steps of the recording are given below:

- 1- The device is activated with the first command and it is ensured that the patient and the project team attendant who will start the shooting are ready in 15 seconds.
- 2- The recording is started with the second command and it is waited for 10 seconds for calibration.
- 3- The eye closure command is given by voice as the third command and the subject is asked to close his eyes. At the same time, the moment when eye closed is enrolled by putting a marker on the recording by recording attendant.
- 4- Opening the eyes command was given 1 min 10 seconds after the third command. After this command, another marker was placed.
- 5- Steps 3 and 4 are repeated for 4 more times.
- 6- The recording process is finished.

In this setup, 1 minute 10 seconds recording time is given for each command and it is aimed to adapt the subject to the command in the closing or opening of the eye in the first 5 seconds. The last 5-second period was removed from the segments by aiming that the eye closed or open shooting times after each command were exactly 1 minute, the shifting of

marker insertion times or eliminating the artifacts that could occur from the voice command in the subjects.

Before recordings, each subject was reminded that they should not move their hands, arms, legs and heads, as this causes noise on EEG signals during recordings. Environment was prepared according to the following:

- 1- It has been paid attention that it is a quiet area to avoid distractions,
- 2- The subjects were provided to sit comfortably,
- 3- It is ensured that the subjects look at a flat wall in such a way that there was no stimulus in front of them to distract their attention when the eye was open,
- 4- The subjects were asked to obey the commands that would require keeping their eyes open or closed during the shooting,
- 5- Experimenter acted in a way that does not make any distracting movement or sound during recordings,

EEG data were recorded with Emotiv EPOC which is a high-quality wireless EEG device. Line noise was removed automatically by the device. Figure 2 shows the EEG recorder.

Channels present in the device are; AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, AF42. Sampling frequency is 128 Hz and a bandpass filter is applied by the device to keep the frequency content between 0.2 and 45Hz.



Figure 2 Eemotive EPOC EEG Device

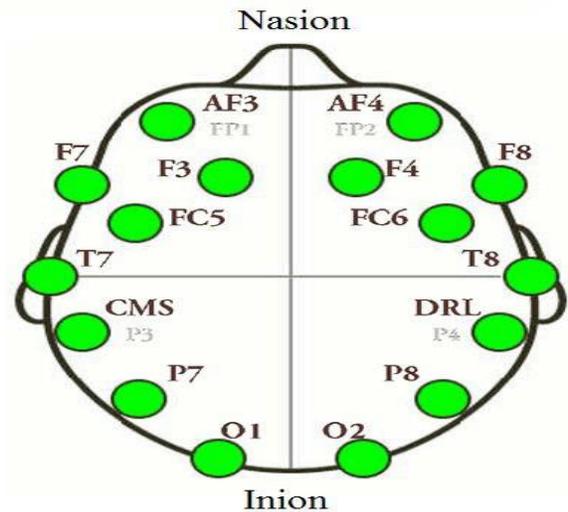


Figure 3 Placement of electrodes according to 10-20 system for EEG recording (top view)

The device is placed on the skull in accordance with the international 10-20 system, and the electrodes are located As shown in Figure 3.

D. Feature Extraction

The collected 7 min and 10 secs raw EEG data were divided into segments. A total of six 1-minute EEG data were obtained, 3 closed eyes and 3 open eyes, for each subject. Features are extracted by employing Continuous Wavelet Transform (CWT). The wavelet model used for CWT is the Morlet wavelet. total energy and energies in theta (3-7Hz), alpha (8-12 Hz), beta (13-29 Hz) and gamma (30-45 Hz) bands obtained for each channel. The total number of features is 14 channels x 5 energy values=70. Two-seconds sliding windows with 50% overlap windows were used for feature extraction.

As a result, a feature matrix of 179x70 was obtained for each eye state (open/closed) from each subject.

E. Classification with kNN Algorithm

In this study, k-Nearest Neighbor (kNN) algorithm is used to classify EEG signals [10,11,12]. K-Nearest Neighborhood (kNN) classification algorithm is a theorem that works according to the distance value between classes and the nearest neighbor value to be selected. According to

this classification algorithm, k points are selected to be closest to the data. The classification process is made according to this selected value. The small value of k helps to collect the sample with the highest similarity rate in the same class while increasing k value can cause different classification results by including dissimilar data in the same class [13]. Therefore, one of the methods used to select the k value is to compare the classification results for various k values, although it takes time. For example, in the study of Özer and Amasya, the classification success obtained when the classification of the data was made according to the 1-nearest neighbor acquired better results than the classifications of 3, 5, 7, 9 and 10-nearest neighboring [14].

III. RESULTS

In this study, the feature matrices obtained from the collected EEG data were created independently from the subject and these matrices were classified according to their neighbor values closest to 1, 3, 5 in the kNN algorithm and the results are given in Table 1.

TABLE I
CLASSIFICATION RESULTS

kNN	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area
1	0.913	0.087	0.913	0.913	0.913	0.913
3	0.897	0.103	0.897	0.897	0.897	0.951
5	0.892	0.108	0.892	0.892	0.892	0.957

EEG recordings collected from 30 patients with clinically depressive disorder for eyes open/closed states were classified with 91.3%, 89.7% and 89.2% accuracy for k = 1, 3 and 5, respectively, according to the kNN classification algorithm. The highest success was obtained for k = 1. As the results show, the EEG signals recorded from patients

with depressive disorder as their eyes were open or closed can be distinguished from each other with a very high accuracy. In future studies, the differentiation accuracies of the EEG signals collected from individuals with clinical depressive disorder and the control group that have the same demographic structure will be examined.

IV. ACKNOWLEDGMENT

This study was supported by the Scientific Research Projects Unit of Adana Alparslan Turkes Science and Technology University. Research Grant No.: 18103031

V. REFERENCES

- [1] T. Erguzel, S. Ozekes, A. Bayram and N. Tarhan, "Classification of major depressive disorder subjects using Pre-rTMS electroencephalography data with support vector machine approach," *2014 Science and Information Conference*, London, 2014, pp. 410-414.
- [2] S. A. Akar, S. Kara, S. Agambayev and V. Bilgiç, "Nonlinear analysis of EEG in major depression with fractal dimensions," *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Milan, 2015, pp. 7410-7413.
- [3] W. Mumtaz, A. S. Malik, S. S. A. Ali and M. A. M. Yasin, "P300 intensities and latencies for major depressive disorder detection," *2015 IEEE International Conference on Signal and Image Processing Applications (ICSIPA)*, Kuala Lumpur, 2015, pp. 542-545.
- [4] H. M. Mallikarjun and H. N. Suresh, "Depression level prediction using EEG signal processing," *2014 International Conference on Contemporary Computing and Informatics (IC3I)*, Mysore, 2014, pp. 928-933.
- [5] A. M. Al-Kaysi, A. Al-Ani, C. K. Loo, M. Breakspear and T. W. Boonstra, "Predicting brain stimulation treatment outcomes of depressed patients through the classification of EEG oscillations," *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Orlando, FL, 2016, pp. 5266-5269.
- [6] Y. Mohan, S. S. Chee, D. K. P. Xin and L. P. Foong, "Artificial neural network for classification of depressive and normal in EEG," *2016 IEEE EMBS Conference on Biomedical Engineering and Sciences (IECBES)*, Kuala Lumpur, 2016, pp. 286-290
- [7] Holden, R. R., Mehta, K., Cunningham, E. J., & McLeod, L. D. (2001). "Development and preliminary validation of a scale of psychache". *Canadian Journal of Behavioural Science/Revue Canadienne Des Sciences Du Comportement*, 33(4), 224.
- [8] Demirkol, M., Güleç, H., Çakmak, S., Namlı, Z., Güleç, M., Güçlü, N., & Tamam, L. (2018). "Reliability and validity study of the Turkish Version of the Psychache Scale". *Anatolian Journal of Psychiatry*, 19(1), 14-20.
- [9] Beck, A. T., & Steer, R. A. (1984). "Internal consistencies of the original and revised Beck Depression Inventory". *Journal of Clinical Psychology*, 40(6), 1365-1367.
- [10] Qu, H. & Gotman, J. A., "Patient-Specific Algorithm for the Detection of Seizure Onset in Long-Term EEG Monitoring: Possible Use as a Warning Device", *IEEE Transactions on Biomedical Engineering*, 44, 115-122p, 1997.
- [11] Oliveria, I., Grigori, O., Guimaraes, N., "EEG Signal Analysis for Silent Visual Reading Classification", *International Journal Of Circuits, Systems And Signal Processing*, Issue 3, Volume 3, 2009:119-126p.
- [12] Murugappan, M., "Human emotion classification using wavelet transform and KNN", *Pattern Analysis and Intelligent Robotics (ICPAIR) International Conference on*, (Volume:1), 28-29 June 2011, 148 - 153p.
- [13] Mitchell, T., 1997, "Machine Learning", McGraw Hill, 1997.
- [14] Özer, Z.B., ve Amasyalı, M.F., "Meta Öğrenme ile KNN Parametre Seçimi", 21. Sinyal İşleme Ve Uygulamaları Konferansı, 2013.

BIOGRAPHIES



Server Göksel ERALDEMİR He received the Bachelor degree in Computer Education and Instructional Technology from the University of OMU, Turkey, in 2002. Then he worked computer teacher

in Reyhanlı high school from 2002 to 2009 and After in late 2009 he started to work as a lecturer at the university. Then he received Masters Degree in electrical and electronics engineering from the University Mustafa Kemal, in 2014. In 2019, I started my doctoral education in the Department of Electrical and Electronics Engineering, Institute of Science and Technology, Adana Alparslan Turkes University of Science and Technology.



Ümit KILIÇ was born in 1992, in Adana, Turkey. He received the BSc degree in Computer Engineering from Çukurova University and the MSc degree in Nanotechnology and Engineering Sciences from Adana Alparslan Turkes Science and Technology University in 2016 and 2019, respectively. He is currently a PhD student in the Department of Electrical-Electronics Engineering in Adana Alparslan Turkes Science and Technology University. He also works as a Research Assistant in the Department of Computer Engineering at the same university. His research interests are Data Mining, Artificial Intelligence, and their applications. He can be reached via ukilic@atu.edu.tr.



Mümine KAYA KELEŞ was born in 1986, in Adana, Turkey. She received the BSc and MSc degrees in Computer Engineering from Çukurova University, Adana, Turkey, in 2009 and 2011, respectively. She received PhD degree in Computer Sciences Division of Electrical and Electronics Engineering from Çukurova University, Adana, Turkey. She is currently Assistant Professor in the Department of Computer Engineering at the Adana Alparslan Turkes Science and Technology University. Her research interests include Data Mining, Distance Education, Text Mining, Plagiarism Detection, Information Retrieval, Text/Document Processing, Web Mining, Database Management Systems, Natural Language Processing, and their applications. Contact her at mkaya@atu.edu.tr. Her ORCID ID is <https://orcid.org/0000-0001-8414-1713>.



Mehmet Emin DEMİRKOL: Mehmet Emin Demirkol graduated from Çukurova University Faculty of Medicine in 2010. Between 2011 and 2015, he received psychiatry specialty training at the same university. He is still working as an instructor at Çukurova University Faculty of Medicine, Department of Psychiatry. He investigated the relationship between suicide attempts and

psychache and alexithymia in patients with schizophrenia. He has studies about bariatric surgery patients and the efficacy of light therapy in depressive patients.



Esen Yildirim, received the B.S. degree in Electrical and Electronics Engineering from Cukurova University, Adana, Turkey, in 1997, and the M.S. and Ph.D. degrees in Electrical Engineering from University of Southern California (USC), Los Angeles, in 2000 and 2006, respectively. She is currently an Associate Professor of Electrical and Electronic Engineering, at Adana Alparslan Türkeş Science and Technology University, Adana, Turkey. Her general research interests include biomedical signal processing, epileptic seizure detection, functional connectivity, learning methods, and emotion recognition from physiological signals.



Lut TAMAM: Professor Doctor Lut Tamam has been working in the field of psychiatry for 25 years and is head of the Department of Psychiatry of Çukurova University Faculty of Medicine for 8 years. He has studies about comorbidity in bipolar disorder, impulse control disorders, attention deficit hyperactivity disorder, light therapy, electroconvulsive therapy, obsessive compulsive disorder, and major depressive disorder. He is also the editor of the journal Current Approaches in Psychiatry.

Effects of Digital Filtering on the Classification Performance of Steady-State Visual Evoked Potential Based Brain-Computer Interfaces

V. ÇETİN, S. OZEKES and H.S. VAROL

Abstract—The electrical activity that occurs during the communication of neurons is recorded by a method called electroencephalography. Brain computer interfaces utilize various electrophysiological sources obtained from different regions of the brain. The electrophysiological source used in this study is the electrical activity seen in the occipital lobes as a result of visual stimuli that flicker at certain frequencies, and is called steady-state visual evoked potential. The main goal in this work is not to try to improve the classification performance but to investigate the effects of different digital filtering algorithms on classification performance. The effects of the high pass and low pass filtering on the classification performance in steady-state visual evoked potential based brain computer interfaces are investigated. As a result of this study, no significant change in the classification performances of designs with only high pass filtering, and high and low pass filtering, has been observed. In addition, it has been observed that only the designs include a high-pass filter implementation give better classification performance in many cases. Consequently, it is concluded that low-pass filtering in steady-state visual evoked potential based brain-computer interfaces does not provide the desired contribution to classification performance.

Index Terms— Brain-Computer Interfaces, Machine Learning, Biomedical Signal Processing, Digital Filters, Intelligent Systems.

I. INTRODUCTION

EVOKED ELECTRICAL signals, caused by a visual stimulus are called visual evoked potentials and are recorded from the occipital and parietal lobes of the brain. Stimuli of frequencies higher than 3.5 Hz are called steady-

state visual evoked potentials (SSVEP) because they produce a quasi-sinusoidal oscillation at the same frequency as the stimulus due to overlapping of excited action potentials [1]. SSVEP is best observed when the stimulus frequency is approximately 15 Hz while a light source that stimulates the retina at any frequency between 3.5 Hz and 75 Hz generates a signal of the same frequency in the visual cortex[2].

Researchers studied on various classification problems using the SSVEP-based BCI systems. Lalor et al. showed that control of computer games can be performed using brain signals with a SSVEP-based BCI system [3]. They applied a binary classification on EEG acquired from O1 and O2 electrodes, and used two stimuli with frequencies of 6 Hz and 25 Hz. They applied a spectral power density (SPD)-based classification, and used the arithmetic mean of the SPD values obtained from O1 and O2 electrodes as features. Kelly et al. performed a left and right classification on EEG obtained from O1 and O2 electrodes using two stimuli at frequencies of 10 Hz and 12 Hz [4]. They applied a SPD-based classification algorithm. Muller-Putz and Pfurtscheller demonstrated that a four-task classification of a biaxial hand prosthesis by using a SSVEP-based BCI that makes it possible the use of BCI based neuro-prosthesis in situations such as spinal cord injury [5]. Prueckl and Guger developed a four-task classification of forward, backward, right and left with a SSVEP-based BCI using four stimuli at 10 Hz, 11 Hz, 12 Hz and 13 Hz frequencies [1]. They used the first and second harmonics obtained from the frequency spectrum of the SSVEP. They acquired the EEG from O1, O2, Oz, PO3, PO4, PO7, PO8 and POz electrodes, and perform a SPD-based classification. Bin et al. designed a SSVEP-based BCI system by performing a canonical correlation based classification on EEG [6]. They acquired EEG from nine channels from occipital and temporal regions of the brain. Luo and Sullivan performed a four-task classification of SSVEP by using only the PO2 electrode [2]. They applied a SPD-based classification, and used four SSVEP stimuli at 9 Hz, 10 Hz, 11 Hz and 12 Hz frequencies. Volosyak developed a SSVEP-based spelling system with a classification of six tasks Based on the five different SSVEP stimuli at frequencies 6.67 Hz, 7.5 Hz, 8.57 Hz, 10 Hz and 12

VOLKAN ÇETİN, is with Department of Computer Engineering Istanbul Arel University, Istanbul, Turkey, (e-mail: volkancetin@arel.edu.tr).

 <https://orcid.org/0000-0003-3864-9246>

SERHAT OZEKES, is with Department of Computer Engineering, Uskudar University, Istanbul, Turkey, (e-mail: serhat.ozekes@uskudar.edu.tr).

 <https://orcid.org/0000-0002-2306-6008>

HÜSEYİN SELÇUK VAROL, is with Department of Electronics and Communications Engineering Doğuş University, Istanbul, Turkey, (e-mail: hsvarol@dogus.edu.tr.com).

 <https://orcid.org/0000-0002-3968-4230>

Manuscript received December 02, 2019; accepted January 26, 2020.

DOI: [10.17694/bajece.654288](https://doi.org/10.17694/bajece.654288)

Hz [7]. Long et al. performed a battery-powered chair control using a hybrid BCI approach [8]. While using the sensorimotor based EEG for the right-left rotation and deceleration of the chair, they used the SSVEP signals obtained from O1, O2 and Oz electrodes for acceleration. Lee et al. controlled a mobile robot using SSVEP stimuli at 13 Hz, 14 Hz and 15 Hz frequencies, corresponding to three different orientations as forward, right and left [9]. Zhang et al. designed an SSVEP-based BCI using canonical correlation-based classification, and stimuli at 6 Hz, 7 Hz, 8 Hz and 9 Hz frequencies. They obtained EEG from the electrodes O1, O2, Oz, P7, P8, P3, P4 and Pz [10].

In general, stimuli that oscillate at frequencies higher than a certain threshold value are called critical vibration frequency (CVF). Frequencies in the range of 50 Hz to 60 Hz are perceived as not vibrating by the visual cortex. This situation results in less eye fatigue in subjects [11]. Sakurada et al. designed a BCI system in which stimuli were used at frequencies above the CVF [11]. The frequencies of 41 Hz, 43 Hz and 45 Hz are used as stimuli that are below CVF, and 61 Hz, 63 Hz and 65 Hz as stimuli above the CVF. They compared the classification results of the system. Although frequencies above the CTF had a slight decline in the classification performance of the BCI, subjects reported less visual fatigue in stimuli at these frequencies. They have shown that high frequency visual stimuli above the CVF can be used in the design of SSVEP-based BCI systems. In previous years several studies has also been conducted to show the connection of the classification performance of EEG and the volume of music, [12] the effect of stimuli frequencies on eyes[11] and the analysis of the performance of different age groups [13]. On the other hand, there is no study has been observed in the literature showing the connection between classification performance and digital filtering algorithms in SSVEP-based BCI applications.

Brain-computer interface (BCI) design usually consists of signal processing, feature extraction and classification stages. In the signal processing stage, it is aimed to improve the signal quality through digital filtering. On the other hand, Widmann et al. have come to the conclusion that digital filtering should not be regarded as an essential step in EEG studies [14]. Vanrullen has reported that data associated with an event related activity can spread over hundreds of milliseconds on EEG due to filtering, and that can cause problems in assessing the timing and dynamics of brain activities [15]. He even advised not to apply low pass filtering when possible. The

question in this case is the acceptability of decreasing classification performance, if any, versus decreasing time complexity of the filtering algorithm, since no digital filtering or only high pass filtering is applied. This study tries to answer this question by comparing the classification performances in cases where certain conventional digital filtering methods are used and not used.

II. MATERIALS AND METHODS

EEG recordings obtained from three subjects were used in this study. The stimulus frequencies, each corresponding to a separate task, are given in Table 1. A four-minute EEG was recorded for each task. Subjects were rested for three minutes between recordings. The user interface of the stimulus program is shown in Figure 1. The program consists of a maximum of 6 stimuli flickering at the desired frequencies. The distances of the flickers to each other and to the edges of the monitor are shown in Figure 1. These distances were measured on a 21.5" monitor when the program was full screen. During the EEG recordings the distance from the monitor to the subjects is set to be approximately 60 cm.

TABLE I
FUNDAMENTAL FREQUENCIES AND HARMONICS OF STIMULI IN HERTZ

Stimuli	1 st Harmonic	2 nd Harmonic	3 rd Harmonic	4 th Harmonic
1	4	8	12	16
2	4.6	9.1	13.7	18.2
3	5.3	10.6	15.9	21.2
4	6.4	12.7	19.1	25.5

A neuro-headset, manufactured by Emotiv, was used for EEG recordings. There are 16 electrodes on the headset, two of which are reference electrodes. The active electrodes in the neuro-headset are in fixed positions as AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, AF4. P3 and P4 are used as reference electrodes. In this study, EEG signals obtained from occipital lobes were required because a SSVEP- based BCI design was analyzed. Therefore, only the data obtained from the channels O1 and O2 are used.

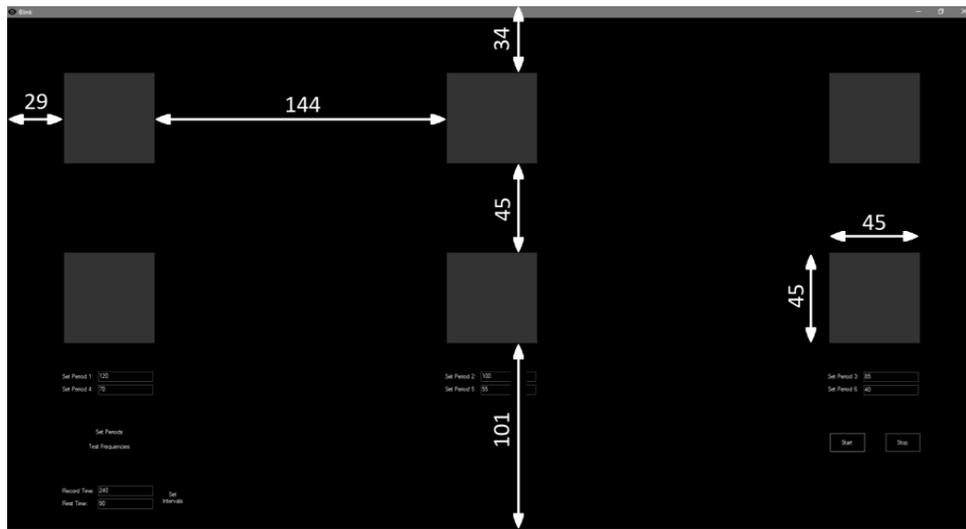


Fig. 1. User Interface of the Stimulus Program, and distances of the flickers to each other and to the edges of the monitor

A. Digital Filter Characteristics

Butterworth infinite impulse response (IIR) filters or finite impulse response (FIR) filters are usually applied in electrophysiology [14]. Generally, two parameters are taken into consideration in IIR or FIR filter design preference. These parameters are system stability and linear phase. IIR filters are not always stable due to feedback. On the other hand, FIR filters are always stable because their design does not include any feedback. In addition, symmetric FIR filters always have a linear phase. When evaluated in terms of system stability and linear phase, the superiority of FIR filters over IIR filters is evident. The advantage of IIR filters over FIR filters is that the desired filter characteristics can be achieved with low-order filters. In this study, FIR filters were applied to avoid possible signal distortions due to non-linear phase. In addition, FIR filter designs, which the filter length is usually relatively longer, are thought to serve better for the purpose of this study. For instance, a 40th-order filter causes more time delay than a 5th-order filter. Thus, the effect of a significant time delay on the classification performance can be better observed by applying a 40th-order filter. Figure 2 and Figure 3 show the frequency (black curve) and phase (grey line) spectra of high-pass (HPF) and low-pass (LPF) filters applied on EEG signals in this study.

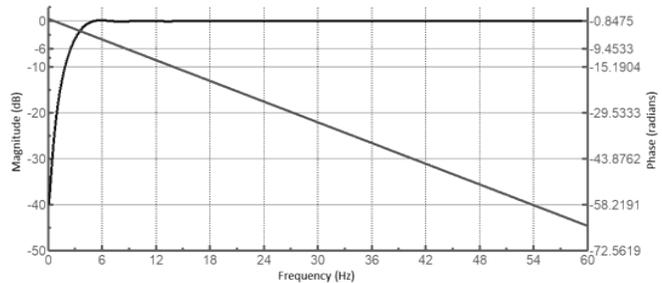


Fig. 2. Frequency and phase spectra of high pass filter applied on EEG data

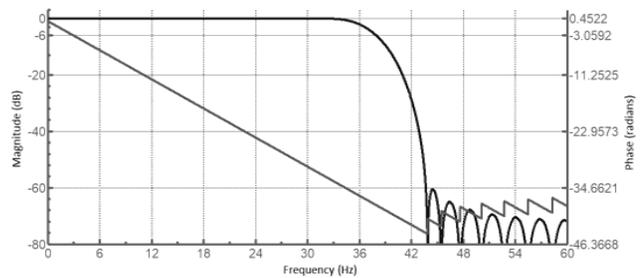


Fig. 3. Frequency and phase spectra of low pass filter applied on EEG data

As one approaches the ideal filter characteristics, the time delay resulting from filtering increases. The time delay of a filter is given by

$$\Delta t = \frac{N-1}{2} T \tag{1}$$

where N is the filter length and T is the sampling period. The HP and LP filters applied in this study have 172 ms and 160 ms time delays respectively.

Filter length is an important parameter, especially when the time windows are relatively shorter. When a signal sampled at 128 Hz is analyzed in two-second windows, each piece of data segment consists of 256 samples. Thus, the HPF with a length of 45 makes the first 22 samples invalid which is 8.6% of the data, as shown in Figure 4(a). This is the edge effect created by the filter on the edges of the data. The edge effect caused by two filters with lengths 79 and 171 can be observed in Figure 4(b) and Figure 4(c), respectively. As one approaches the ideal filter characteristics, the edge effect due to the increase of the filter length causes a significant data loss which cannot be accepted especially in real time applications.

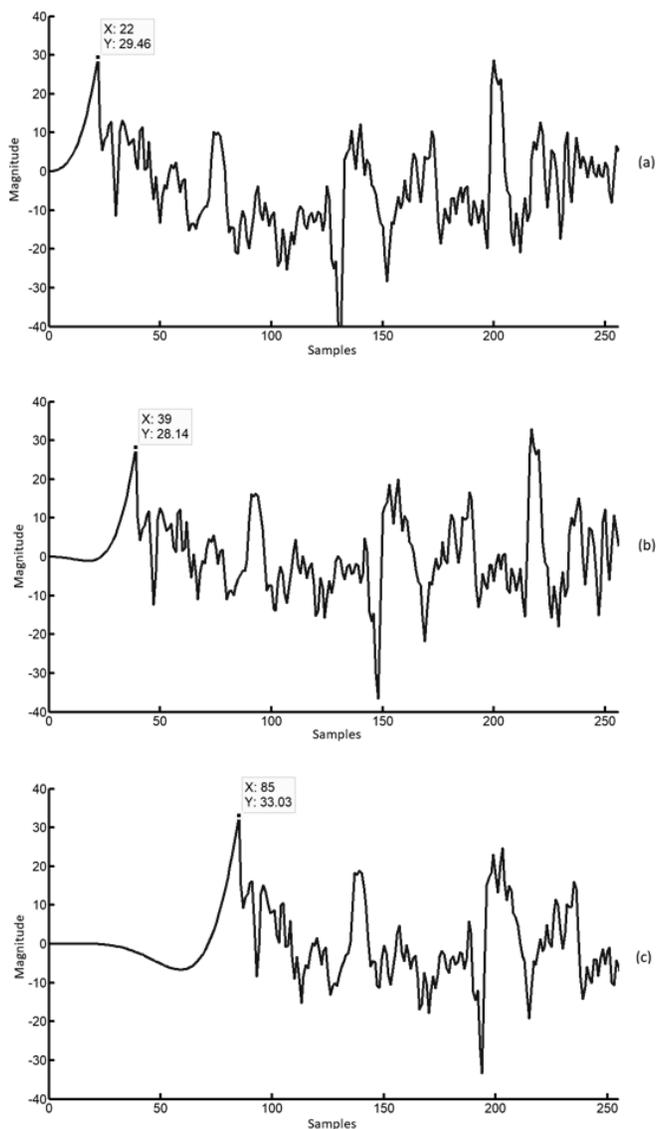


Fig. 4. Edge effect due to filters with lengths of (a) 45, (b) 79 and (c) 171, respectively

B. Classification

In general, it is observed that the most widely used classification methods in the BCI literature are artificial neural

networks (ANN), support vector machines (SVM) and Naive Bayes (NB) [16,17]. In this study, a three layered artificial neural network model was used for classification.

SSVEP-based BCI users are best at modulating the spectral properties of the stimuli. Therefore, the spectral properties of EEG signals are mostly used as differentiating features for classification in SSVEP-based BCI applications [1,18-23]. For this reason the feature vector consists of the total and relative band power values of the first and second harmonics obtained from the frequency spectra of the SSVEPs. In this study, binary, ternary and quaternary classifications were performed. Three-layered perceptron network, trained by back-propagation algorithm, was used in classification. The input and the output layers contain neurons as many as the number of features and number of classes, respectively. Number of neurons in hidden layer is set to be the mean of the number of neurons in the input and the output layers. The learning rate of the network is 0.3 and the training iteration is 40. A 10-fold cross-validation model was used to evaluate the classification results. Each fold is composed of random samples with equal class distribution.

III. RESULTS

Four different cases were investigated in this study. These are cases which no digital filtering is applied, only HPF is applied, only LPF is applied, and HPF and LPF are applied sequentially. The applied HPF and LPF characteristics are given in Table 2.

TABLE II
HPF AND LPF CHARACTERISTICS APPLIED IN FILTERING EEG SIGNALS

	HPF	LPF
Attenuation	37 dB	60.5 dB
Passband Ripple	0.33 dB	0.1 dB
Transition Bandwidth	4.6 Hz	12 Hz
6 dB Point	2.48 Hz	38 Hz
3 dB Point	3.22 Hz	36.7 Hz
Filter Length	45	42

The accuracy percentages of the binary, ternary and quaternary classifications using 2 and 4 second window lengths are given in Table 3. There are two important conclusions from Table 3. First, digital HPF increases classification performance by an average of 9.38% in all classifications of all subjects. The fact that the digital HPF improves classification performance in all cases, independent of time window, subject and number of classes, suggests that the electronic HPF on neuro-headset alone does not show adequate filtering performance for a BCI application. On the

other hand, it is observed that the digital LPF does not have a significant effect on the classification performance and in most cases it has a negative effect. This can be explained by the fact that the electronic LPF in the neuro-headset is capable of providing adequate classification performance. Another explanation may be that the time delay and edge effect due to digital LPF have too much effect, while the low energy of high frequency noises has a less negative effect on classification performance.

TABLE III

THE ACCURACY PERCENTAGES OF THE BINARY, TERNARY AND QUATERNARY CLASSIFICATIONS USING 2 AND 4 SECOND WINDOW LENGTHS

		2s			4s		
		2Tasks	3Tasks	4Tasks	2 Tasks	3 Tasks	4 Tasks
Subject 1	No filtering	68.75	55.28	46.04	75.83	59.44	46.67
	Only LPF	66.25	56.94	45.00	69.17	58.33	51.25
	Only HPF	78.75	62.22	52.71	82.50	64.44	59.58
	LPF + HPF	78.75	57.78	49.17	78.33	67.22	60.83
Subject 2	No filtering	92.08	70.83	53.54	90.00	71.11	57.92
	Only LPF	90.00	67.78	51.67	85.00	61.67	61.67
	Only HPF	96.67	85.83	66.46	94.17	82.22	55.42
	LPF + HPF	97.50	85.00	65.00	93.33	85.56	59.58
Subject 3	No filtering	71.25	56.11	51.25	81.67	64.44	57.08
	Only LPF	73.33	54.17	51.46	79.17	59.44	55.83
	Only HPF	82.92	69.22	67.92	87.50	75.00	74.58
	LPF + HPF	80.83	70.56	67.29	87.50	73.89	70.83
Mean	No filtering	77.36	60.74	50.28	82.50	65.00	53.89
	Only LPF	76.53	59.63	49.38	77.78	59.81	56.25
	Only HPF	86.11	72.42	62.36	88.06	73.89	63.19
	LPF + HPF	85.69	71.11	60.49	86.39	75.56	63.75

The second important conclusion is understood when comparing analyzes that includes applying only HPF and that includes applying HPF and LPF. Table 4 gives the average classification performances obtained from the aforementioned analyzes. It is seen that HPF and LPF together affected the classification performance at a range of 1.67% to -1.87%. Despite LPF-caused time delay, the classification performance increases up to 1.67% at most. In addition, it has been observed that the case which includes applying HPF and LPF together decreases the classification performance between 0.42% and 1.87% in certain cases. Results obtained from this study verify that the digital filtering is not an absolute step in BCI analysis as well as concluded by Widmann et al. for EEG analysis. It

also verifies that LPF should not be applied unless there is a significant necessity as Vanrullen proposes.

TABLE IV
COMPARISON OF CLASSIFICATION PERFORMANCES OF BCI DESIGNS

	2s			4s		
	2 Tasks	3 Tasks	4 Tasks	2 Tasks	3 Tasks	4 Tasks
Only HPF	86.11	72.42	62.36	88.06	73.89	63.19
HPF + LPF	85.69	71.11	60.49	86.39	75.56	63.75
Difference	-0.42	-1.31	-1.87	-1.67	1.67	0.56

IV. CONCLUSION

In this study, the effect of different digital filtering strategies on the classification performance of BCI applications is investigated. We have experimentally confirmed the conclusions and the suggestions with the adverse effects of the digital filtering in EEG analysis mentioned in previous publications. This study shows that low-pass filtering in BCI applications does not provide the desired contribution to classification performance. On the other hand when evaluating the results obtained from this study the characteristics of the EEG device and digital filters should not be overlooked. The neuro-headset has an HPF at cut-off frequency of 0.15 Hz and a LPF at cut-off frequency of 43 Hz. Different EEG recorders could probably contain electronic filters with different characteristics or no electronic filters at all. Therefore there should be no definite conclusion that LPF should not be applied in BCI applications. Instead, the effect of digital filters on the classification performance should be investigated and a digital filtering strategy should be determined according to the results in BCI applications. Another conclusion can also be deduced that the electronic filters contained in the EEG recording devices should not be relied upon. Although the neuro-headset used in this study have a HPF at a cut-off frequency of 0.15 Hz, an additional digital HPF has improved the classification results.

References

- [1] R. Prueckl, C. Guger, "A Brain-Computer Interface Based on Steady State Visual Evoked Potentials for Controlling a Robot." 10th International Work-Conference on Artificial Neural Networks, vol.10, no. 12 June, Salamanca, Spain. 2009. doi:https://doi.org/10.1007/978-3-642-02478-8_86
- [2] A.Luo, T.J. Sullivan, "A User-Friendly SSVEP-Based Brain-Computer Interface using a Time-Domain Classifier." Journal of Neural Engineering, vol. 7, no. 2, pp. 026010, 2010. doi:10.1088/1741-2560/7/2/026010
- [3] E.C. Lalor, S.P. Kelly, C. Finucane, R. Burke, R. Smith, R.B. Reilly, G. McDarby, "Steady-State VEP-Based Brain-Computer Interface Control in an Immersive 3D Gaming Environment." EURASIP Journal on Applied Signal Processing, vol. 2005, no. 19, pp. 3156-3164, 2005. doi: <https://doi.org/10.1155/ASP.2005.3156>
- [4] S.P. Kelly, E.C. Lalor, C. Finucane, G. McDarby, R.B. Reilly, "Visual Spatial Attention Control in an Independent Brain-Computer Interface."

- IEEE Transactions on Biomedical Engineering, vol. 52 np. 9, pp. 1588-1596, 2005. doi: 10.1109/TBME.2005.851510
- [5] G.R. Muller-Putz, G. Pfurtscheller, "Control of an Electrical Prosthesis with an SSVEP-Based BCI." IEEE Transactions on Biomedical Engineering, vol. 55, no. 1, pp. 361-364, 2008. doi: 10.1109/TBME.2007.897815
- [6] G. Bin, X. Gao, Z. Yan, B. Hong, S. Gao, "An Online Multi-Channel SSVEP-Based Brain-Computer Interface using a Canonical Correlation Analysis Method. Journal of Neural Engineering", vol. 6, no. 4, pp. 046002, 2009. doi: 10.1088/1741-2560/6/4/046002
- [7] I. Volosyak, "SSVEP-Based Bremen-BCI Interface – Boosting Information Transfer Rates." Journal of Neural Engineering, vol. 8, no.3, pp. 036020, 2011. doi: 10.1088/1741-2560/8/3/036020
- [8] J. Long, Y. Li, H. Wang, T. Yu., J. Pan, F. Li, "A Hybrid Brain Computer Interface to Control the Direction and Speed of a Simulated or Real Wheelchair." IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 20, no. 5, pp. 720-729, 2012. doi: 10.1109/TNSRE.2012.2197221
- [9] P. Lee, H. Chang, T. Hsieh, H. Deng, C. Sun, "A Brain-Wave- Actuated Small Robot Car using Ensemble Empirical Mode Decomposition-Based Approach." IEEE Transactions on Systems, Man, and Cybernetics – Part A: Systems And Humans, vol. 42, no. 5, pp. 1053-1064, 2012. doi:10.1109/TSMCA.2012.2187184
- [10] Y. Zhang, G. Zhou, J. Jin, X. Wang, A. Cichocki, "SSVEP Recognition using Common Feature Analysis in Brain-Computer Interface." Journal of Neuroscience Methods, vol. 244, pp. 8-15, 2015. doi: 10.1016/j.jneumeth.2014.03.012
- [11] T. Sakurada, T. Kawase, T. Komatsu, K. Kansaku, "Use of High-Frequency Visual Stimuli Above the Critical Flicker Frequency in a SSVEP Based BMI." Clinical Neurophysiology, vol. 126, no. 10, pp. 1972-1978, 2015. doi: 10.1016/j.clinph.2014.12.010
- [12] K. Hasan, S. Hossain, T.K. Ghosh, M. Ahmad, "A SSVEP Based EEG Signal Analysis to Discriminate the Effects of Music Levels on Executional Attention." American Journal of Bioscience and Bioengineering, vol. 3, no. 3-1, pp. 27-33, 2015. doi: 10.11648/j.bio.s.2015030301.15
- [13] F. Gembler, P. Stawicki, I. Volosyak, "A Comparison of SSVEP-Based BCI-Performance Between Different Age Groups." 13th International Work-Conference on Artificial Neural Networks, 10-12 June, Palma de Mallorca, Spain, 2015. doi: https://doi.org/10.1007/978-3-319-19258-1_6
- [14] A. Widmann, E. Schröger, B. Maess, "Digital Filter Design for Electrophysiological Data - A Practical Approach. Journal of Neuroscience Methods", vol.250, pp. 34-46, 2014. doi: 10.1016/j.jneumeth.2014.08.002
- [15] R. Vanrullen, "Four Common Conceptual Fallacies in Mapping the Time Course of Recognition." Frontiers in Psychology, vol. 2, Article365, 2011. doi: 10.3389/fpsyg.2011.00365
- [16] L.F. Nicolas-Alonso, J. Gomez-Gil, "Brain Computer Interfaces, a Review." Sensors, vol. 12, no. 2, pp. 1211-1279, 2012. doi: 10.3390/s120201211
- [17] F. Lotte, M. Congedo, A. Lecuyer, F. Lamarche, B. Arnaldi, "A Review of Classification Algorithms for EEG-Based Brain-Computer Interfaces." Journal of Neural Engineering, vol. 4 no. 2, R1, 2007. doi: 10.1088/1741-2552/aab2f2
- [18] I. Volosyak, "SSVEP-Based Bremen-BCI Interface - Boosting Information Transfer Rates." Journal of Neural Engineering, vol. 8, no. 3, pp. 036020, 2011. doi: 10.1088/1741-2560/8/3/036020
- [19] E.C. Lalor, S.P. Kelly, C. Finucane, R. Burke, R. Smith, R.B. Reilly, G. McDarby, "Steady-State VEP-Based Brain-Computer Interface Control in an Immersive 3D Gaming Environment." EURASIP Journal on Applied Signal Processing, vol. 2005 no. 19, pp. 3156-3164, 2005. doi:10.1155/ASP.2005.3156
- [20] P. Lee, H. Chang, T. Hsieh, H. Deng, C. Sun, "A Brain-Wave-Actuated Small Robot Car using Ensemble Empirical Mode Decomposition-Based Approach." IEEE Transactions on Systems, Man, and Cybernetics -PartA: Systems And Humans, vol. 42, no. 5, pp. 1053-1064, 2012. doi: 10.1109/TSMCA.2012.2187184
- [21] J.R. Wolpaw, E.W. Wolpaw, "Brain-Computer Interfaces Principles and Practice, Oxford University Press, Inc.", New York, USA, 2012. doi: 10.1093/acprof:oso/9780195388855.001.0001
- [22] Y. Wang, R. Wang, X. Gao, B. Hong, S. Gao, "A Practical VEP- Based Brain-Computer Interface." IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 14, no. 2, 234-239, 2006. doi: 10.1109/TNSRE.2006.875576
- [23] H. Gollee, I. Volosyak, A.J. McLachlan, K.J. Hunt, A. Graser, "An SSVEP-Based Brain-Computer Interface for the Control of Functional Electrical Stimulation." IEEE Transactions on Biomedical Engineering, vol. 57, no. 8, pp. 1847-1855, 2010. doi:10.1109/TBME.2010.2043432

BIOGRAPHIES



VOLKAN ÇETİN was born in Burdur in 1984. He is a Ph.D. candidate in Institute of Pure and Applied Sciences of Marmara University. He gives lectures in Computer Engineering Department of Istanbul Arel University and Computer Engineering Department of Istanbul Sabahattin Zaim University.



SERHAT OZEKES was born in Washington D.C., USA in 1978. He received his B.S., M.S. and Ph.D. degrees from Marmara University, Computer and Control Education Department in 2000, 2002 and 2006, respectively. He has published more than 25 papers in various international journals and conference proceedings on data mining, image processing and EEG signal processing.



HÜSEYİN SELÇUK VAROL was born in Ankara in 1948. He received his Bachelor's and Master's degrees from the Middle East Technical University Physics Department in 1969 and 1972 respectively. He completed his PhD in Physical Electronics at Surrey University in England, and worked in many companies, universities and research institutions. He has been working at the Department of Electronics and Communications Engineering at Doğuş University, where he has been working since 2015. He has over 80 international and national publications, and his current area of scientific interest is biomedical image and signal processing.

Publication Ethics

The journal publishes original papers in the extensive field of Electrical-electronics and Computer engineering. To that end, it is essential that all who participate in producing the journal conduct themselves as authors, reviewers, editors, and publishers in accord with the highest level of professional ethics and standards. Plagiarism or self-plagiarism constitutes unethical scientific behavior and is never acceptable.

By submitting a manuscript to this journal, each author explicitly confirms that the manuscript meets the highest ethical standards for authors and coauthors

The undersigned hereby assign(s) to *Balkan Journal of Electrical & Computer Engineering* (BAJECE) copyright ownership in the above Paper, effective if and when the Paper is accepted for publication by BAJECE and to the extent transferable under applicable national law. This assignment gives BAJECE the right to register copyright to the Paper in its name as claimant and to publish the Paper in any print or electronic medium.

Authors, or their employers in the case of works made for hire, retain the following rights:

1. All proprietary rights other than copyright, including patent rights.
2. The right to make and distribute copies of the Paper for internal purposes.
3. The right to use the material for lecture or classroom purposes.
4. The right to prepare derivative publications based on the Paper, including books or book chapters, journal papers, and magazine articles, provided that publication of a derivative work occurs subsequent to the official date of publication by BAJECE.
5. The right to post an author-prepared version or an official version (preferred version) of the published paper on an internal or external server controlled exclusively by the author/employer, provided that (a) such posting is noncommercial in nature and the paper is made available to users without charge; (b) a copyright notice and full citation appear with the paper, and (c) a link to BAJECE's official online version of the abstract is provided using the DOI (Document Object Identifier) link.



ISSN: 2147- 284X
Year: January 2020
Volume: 8
Issue: 1

CONTENTS

A. Özmen, B. Tander, H. Şenol; Performance of Cellular Neural Network Based Channel Equalizers	1-6
E. Dursun, S. Varbak Nese, B. Kilic; Green building certification of urban public railway transport systems for sustainable cities,	7-15
N Gedik; A New Feature Extraction Approach Using Contourlet Transform and T-Test Statistics for Mammogram Classification,.....	16-20
C. Catal, A. Kassahun, H. Jan Hoving; Improving Farm Management Information Systems with Data Mining,	21-30
M. Gökdağ, O. Gülbudak; Model Predictive Control of an Indirect Matrix Converter with Active Damping Capability,	31-39
F. Erken; The Impact of the Government's Incentives on Increasing Investment in Turkey's Solar Photovoltaic Power Plants,	40-49
İ. F. Kılınçer, F. Ertam, A. Şengür; Automated Fake Access Point Attack Detection and Prevention System with IoT Devices,.....	50-56
H. E. Kızıloz; On Base Station Localization in Wireless Sensor Networks,.....	57-61
T. C. Karalar; Desinging Analog Mixed Signal Circuits Using Graphene Nano Ribbon Field Effect Transistors,.....	62-66
C. Rahebi, M. Al-Jumaili; A New Bat Optimization Algorithm to Solve EPD Problem Solving with Transmission Loss,.....	67-72
Z. Doğan, R. Selçuk; A Diagnosis of Stator Winding Fault Based on Empirical Mode Decomposition in PMSMs,.....	73-80
G. Kayhan, E. Ergün; Medicinal and Aromatic Plants Identification Using Machine Learning Methods,	81-87
L. Gökrem, M. S. Can, S. Aydın; Hexapod Robot Design and Performance Comparison of Fuzzy and PID Control Methods,.....	88-97
H. Karakaya, İ. E. Şen; Phase Changing Material Usage to Increase the Efficiency of Photovoltaic Panels,	98-102
S. G. Eraldemir, Ü. Kiliç, M. Keleş, M. E. Demirkol, E. Yildirim, L. Tamam; Classification of EEG Signals in Depressed Patients,	103-107
V. Çetin, S. Özkes, H. S. Varol; Effects of Digital Filtering on the Classification Performance of Steady-State Visual Evoked Potential Based Brain-Computer Interfaces,.....	108-113

BALKAN JOURNAL OF ELECTRICAL & COMPUTER ENGINEERING

(An International Peer Reviewed, Indexed and Open Access Journal)

Contact

Batman University
Department of Electrical-Electronics Engineering
Bati Raman Campus Batman-Turkey

Web: <http://dergipark.gov.tr/bajece>
<http://www.bajece.com>
e-mail: bajece@hotmail.com

