SELCUK UNIVERSITY
FACULTY OF TECHNOLOGY

ICENTE'20

INTERNATIONAL CONFERENCE ON ENGINEERING TECHNOLOGIES

November 19-21, 2020
Konya/TURKEY

PROCEEDINGS

Editor
Prof. Dr Sakir TASDEMIR

E-ISBN: 978-625-44427-4-2

International Conference on Engineering Technologies

4th International Conference, ICENTE
Konya, Turkey, November 19-21, 2020

Proceedings

Editor
Sakir TASDEMIR

International Conference on Engineering Technologies, **ICENTE'20**
Konya, Turkey, November 19-21, 2020

International Conference on Engineering Technologies

4th International Conference, ICENTE
Konya, Turkey, November 19-21, 2020

Proceedings

Editors
Sakir TASDEMIR

E-ISBN: 978-625-44427-4-2

SN Bilgi Teknolojileri
Tel: 0.332 323 07 39
www.snbt.com.tr

November – 2020

EDITORS :

Prof. Dr. Sakir TASDEMIR
Selcuk University, Turkey
Depertmant of Computer Engineering, Faculty of Technology
Alaeddin Keykubat Campus, 42031, Konya, TURKEY
stasdemir@selcuk.edu.tr

ASSISTANT EDITORS :

Ilker Ali OZKAN
Selcuk University, Turkey
Depertmant of Computer Engineering, Faculty of Technology
Alaeddin Keykubat Campus, 42031, Konya, TURKEY
ilkerozkan@selcuk.edu.tr

Murat KOKLU
Selcuk University, Turkey
Depertmant of Computer Engineering, Faculty of Technology
Alaeddin Keykubat Campus, 42031, Konya, TURKEY
mkoklu@selcuk.edu.tr

PREFACE

International Conference on Engineering Technologies (ICENTE'20) was organized in Konya, Turkey on 19-21 November 2020.

The main objective of ICENTE'20 is to present the latest research and results of scientists related to Biomedical, Computer, Electrics & Electronics, Mechanical, Mechatronics, Metallurgy & Materials and Civil Engineering fields. This conference provides opportunities for the delegates from different areas in order to exchange new ideas and application experiences, to establish business or research relations and to find global partners face to face for future collaborations.

All paper submissions have been double blind and peer reviewed and evaluated based on originality, technical and/or research content/depth, correctness, relevance to conference, contributions, and readability. Selected papers presented in the conference that match with the topics of the journals will be published in the following journals:

- Gazi Journal of Engineering Sciences (GJES)
- International Journal of Applied Mathematics, Electronics and Computers (IJAMEC)
- International Journal of Automotive Engineering and Technologies (IJAET)
- International Journal of Energy Applications and Technology (IJEAT)
- MANAS Journal of Engineering (MJEN)
- Selcuk-Teknik Journal (SUTOD)

At this conference, there are 185 paper submissions. Each paper proposal was evaluated by two reviewers. And finally, 154 papers were presented at the conference from 9 different countries (Brazil, Spain, Iranian, Kyrgyzstan, Macedonia, Pakistan, Poland, Serbia, Turkey) with 98 local and foreign universities and organizations participating,

In particular, to Selcuk University Rector Prof. Dr. Metin AKSOY; we would like to thank the conference scientific committee, session chairs, invited speakers, referees, technical team, participants, and all our colleagues who have contributed. They have made a crucial contribution to the success of this conference. Our thanks also go to our colleagues in our conference office.

Prof. Dr. Sakir TASDEMIR
Editor

PROGRAMME COMMITTEES

HONORARY CHAIR :

Metin Aksoy, Rector of Selcuk University, Turkey

GENERAL CHAIRS :

Sakir Tasdemir, Selcuk University, Turkey

CO-CHAIRS :

Ilker Ali Ozkan, Selcuk University, Turkey
Ismail Saritas, Selcuk University, Turkey
Lilia Georgieva, Heriot Watt University, United Kingdom
Murat Koklu, Selcuk University, Turkey
Silyan Sibinov Arsov, Rousse University, Bulgaria

SECTION EDITORS :

Adem Golcuk, Fen Bilimleri, Turkey
Fatih Basciftci, Selcuk University, Turkey
Ismail Saritas, Selcuk University, Turkey
Mehmet Cunkas, Selcuk University, Turkey
Mustafa Acarer, Selcuk University, Turkey
Mustafa Altin, Konya Technical University, Turkey
Murat Ciniviz, Selcuk University, Turkey

INTERNATIONAL ADVISORY BOARD :

Abdulkadir Saday, Selcuk University, Turkey
Abdullah Erdal Tumer, Kirgizistan Turkiye Manas University, Kyrgyzstan
Ahmet Fenercioglu, Gaziosmanpasa University, Turkey
Ahmet Yonetken, Afyon Kocatepe University, Turkey
Ahmet Afsin Kulaksiz, Konya Technical University, Turkey
Alexander Sudnitson, Tallinn University of Technology, Estonia
Ali Yasar, Selcuk University, Turkey
Alina Ivan Dramogir, Gheorghe Asachitechnical University of Iasi, Romania
Almoataz Youssef Abdelaziz, Ain Shams University, Egypt
Amar Ramdane Cherif, University of Versailles, France
Anca Loana Andreescu, Academy of Economic Studies, Bulgaria
Anne Villems, University of Tartu, Estonia
Antonella Reitano, University of Calabria, Italy
Antonio Mendes, Universidade De Coimbra, Portugal
Arif Gok, Amasya Teknoloji Faculty, Turkey
Aristomenis Antoniadis, Technical University of Crete, Greece
Artan Luma, South East European University, MACEDONIA
Bahattin Karakaya, Istanbul University, Turkey
Biagio Lenzitti, University of Palermo, Italy
Binod Kumar, Jspm Jayawant Institute of Computer Applications Pune, India
Boris Akanaev, Kazak National University, Kazakhstan
Domenico Tegolo, Universita Degli Studi Di Palermo, Italy
Eisha Akanksha, Mvj College of Engineering, India
Elinda Kajo Mece, Polytechnic University of Tirana, Romania
Engin Ozdemir, Kocaeli University, Turkey
Erol Turkes, Kirklareli University, Turkey
Ertugrul Durak, Suleyman Demirel University, Turkey
Gabriel Luna Sandoval, Sonora State University, Mexico
Hamit Saruhan, Duzce University, Turkey
Hamza Bensouilah, Laboratoire de Mecanique et Structures, Algeria
Hasan Gokkaya, Karabuk University, Turkey
Howard Duncan, Dublin City University, Ireland

ORGANIZING COMMITTEE :

Adem Golcuk, Fen Bilimleri, Turkey
Ayhan Erol, Afyon Kocatepe University, Turkey
Angel Smrikarov, Rousse University, Bulgaria
Fatih Basciftci, Selcuk University, Turkey
Ismail Saritas, Selcuk University, Turkey
Lilia Georgieva, Heriot Watt University, United Kingdom
Mehmet Cunkas, Selcuk University, Turkey
Murat Ciniviz, Selcuk University, Turkey
Mustafa Acarer, Selcuk University, Turkey
Polyxeni Arapi, Technical University of Greece, Greece
Silyan Sibinov Arsov, Rousse University, Bulgaria

TECHNICAL COMMITTEE :

Esra Kaya, Selcuk University, Turkey
Burak Tezcan, Selcuk University, Turkey
Musa Doğan, Selcuk University, Turkey
Abdulkadir Saday, Selcuk University, Turkey

CONTENTS

Colon Automation System

E. TELLİ[1] and E. ŞEHİRLİ[2]

[1] Karabuk University, Karabük/Turkey, emrahtelli@yahoo.com
[2] Karabuk University, Karabük/Turkey, eftalsehirli@karabuk.edu.tr

Abstract **- Automation systems are designed to supply very high precision and accurate results in almost every work. Therefore, mechanization and computers have joined more and more in our lives. With this perspective, creating an automation system to solve problem of healthcare facilities by the respect of treatments plays a key role to make people's lives easier. In this paper, a robust colon automation system (CAS) is proposed. CAS automatedly detects colon in uploaded computed tomography (CT) colonography images whose extension is Digital Imaging and Communication in Medicine (DICOM). CAS views detected colons based on image reconstruction as a three-dimensional (3D) volume image. CAS gives parametric information to surgeon after processing images. Moreover, it enables surgeon to zoom, measure, rotate the viewing from different aspects without touching any computer parts thanks to Leap Motion controller for preoperative and intraoperative medical purpose. CAS remotely controls surgical clamps, scissors, and forceps during surgery by using Leap Motion controller. For this purpose, a robotic arm that represents forceps behavior is controlled by gesture motions. All information about patients like disease history, surgeries are saved in a database developed in this study. For further use, clinicians' information is kept for both use of system and the possibility of investigation under any bad situations.**

Keywords **- 3D imaging, Colon automation system, Image processing, Leap Motion controller, Robotic arm.**

I. Introduction

WILHELM RONTGEN discovered X-ray in 1895. He saw that X-ray could pass through human body parts with different absorbance ratio and create different contrasts when it exposed to photographic film. Thus, that formed projection radiography (common chest X-ray) is the fundamental of medical imaging [1,2]. In the 1960's Godfrey N. Hounsfield and Allan M. Cormack came up with a new idea for medical imaging. The idea was taking a lot of projections from different angles, combining them by using computers instead of taking single projection and losing a lot of medical data based on shadow way of imaging. That idea created CT and medical imaging has shown a huge progress every year with the advent of new methods [3]. Afterwards, with the developing computerization, 3D imaging has become a better way to view body parts. A developed 3D image serves as faster and accurate opportunity in clinical diagnosis. Before the development of three-dimensional model, it has been a difficulty to view medical images from different perspectives. However, with the development of image processing techniques and medical imaging devices, it is expected almost in any case.

The colon is about 150 cm of the gastrointestinal tract between the ileocecal valve and the rectosigmoid corner. The cecum is respectively continuing with the ascending colon, transverse colon, descending colon, and sigmoid colon then joining with the rectum in the rectosigmoid corner. The floors of the colon wall are mucous membrane, submucosa, inner circular muscle, outer longitudinal muscle, and serosa [4].

Robotics is a rapidly growing sector with the invention of robotic systems. Robotic systems aid people in wide range of working areas such as boring repetitive works, tough tasks, dirty tasks, or tasks that needs absolute precision. These areas include military services, car production, space exploration and healthcare and so on. According to sector, they have different shapes, abilities, hardware, and software designs.

In the literature, there are studies related to 3D reconstruction image, colon detection and robotics arm control. Al-Shayea and Al-Ani [5] explained how to make efficient 3D reconstruction and its methods. Angelopoulou et al. [6] worked on 3D surface reconstruction of sensitive high-resolution ventricular MRI slice images. Wang and Pu [7] worked on their studies to reconstruct CT images with VTK and Visual Studio. Datino et al. [8] worked on checking robotic remote catheter system safety and compare it with manual case. Stoianovici [9] worked on the fact that showing how industrial revolution and improving technology are effective in robotic surgery. Camarillo et al. [10] examined background of robotics in surgery and look situation today and possibilities in future.

With this study, robotics arm hardware controlled by Leap Motion controller, image processing to automatedly detect colon on CT colonography images, reconstruction models to view 3D colon, calculation of parameters related to colon and database system to keep information of patients and surgery progress are combined under a software program named as CAS. It is supplied touchless hardware control on 3D view of colon for surgical purpose and robotic arm. As it is seen from the literature view, studies are performed without combining these features in an automation system.

II. Materials and Methods

MATLAB 2018a program, Visual Studio 2017 program, SQL Server management studio 2014, 3D printed robotic arm kit, Leap Motion controller, Arduino Nano, Lattepanda single board computer and 6524 anonym CT colonography DICOM images belonging to 7 patients provided by Cancer Imaging

Archieve have been used in this study. One sample of image is shown in Figure 1. MATLAB program has been utilized to detect colon from CT colonography images and make 3D imaging. A windows form application has been developed under visual studio using C3 programming language. One sample view of application is shown in Figure 2. Patient information and surgery processes have been stored in developed database under SQL Server. Entity-Relationship (ER) diagram of the database is shown in Figure 3. 3D printed robotics arm kit shown in Figure 4 has been used to supply touchless hardware control. Leap Motion controller has been utilized to detect finger and hand movement and manage robotics arm kits. Internal structure of Leap Motion controller is shown in Figure 5. Arduino Nano has been used to control and manage servo motors of robotic arm. Lattepanda single board computer shown in Figure 6 has been utilized to combine under one platform and connect all hardware and software used in this study.

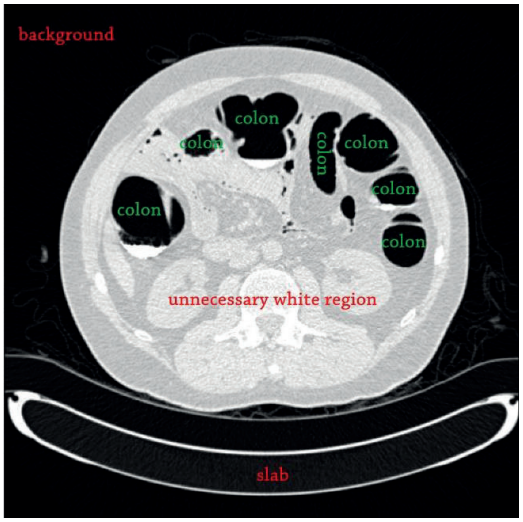


Figure 1: One sample of CT colonography DICOM image.

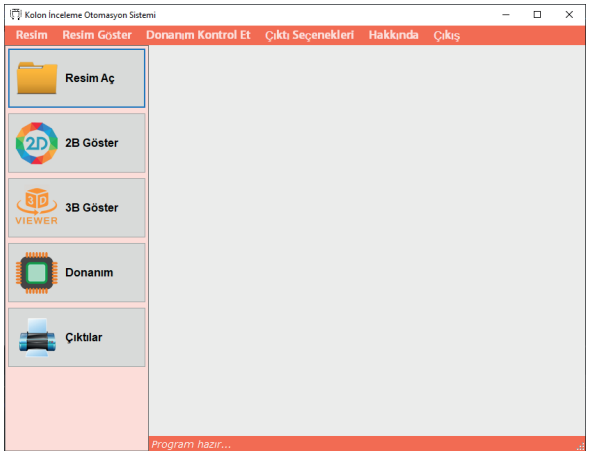


Figure 2: One sample view of colon automation system software.

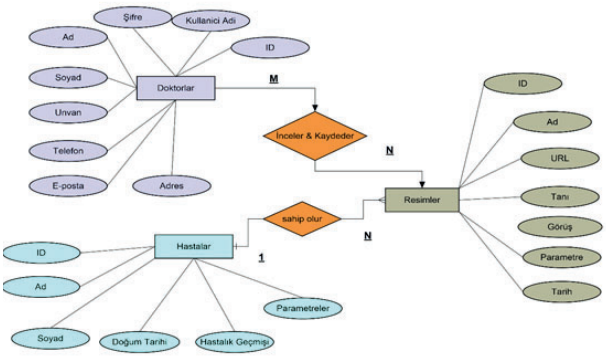


Figure 3: ER diagram of database design.

Figure 4: 3D printed robotics arm kit.

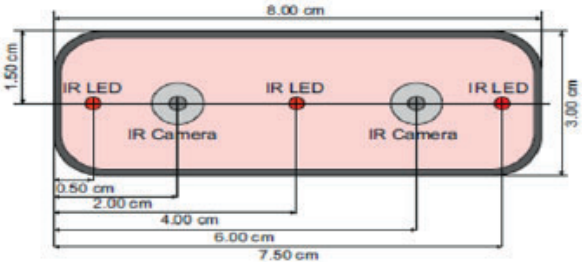


Figure 5: Internal structure of Leap Motion controller [11].

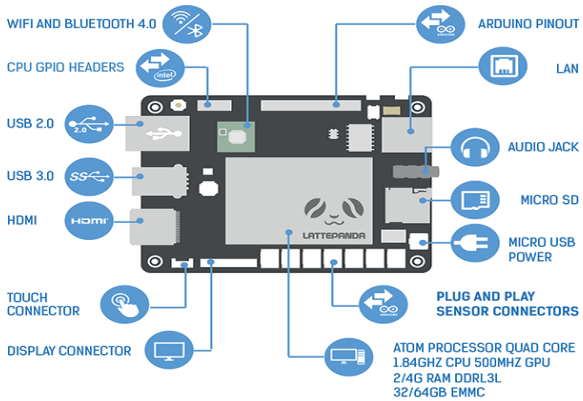


Figure 6: Lattepanda single board computer [12].

The block diagram of CAS is shown in Figure 7 below.

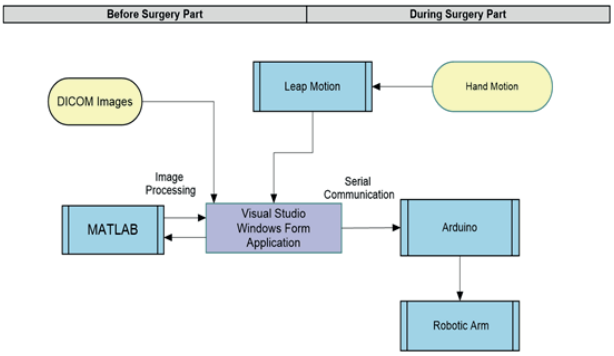


Figure 7: The block diagram of CAS.

A windows form application has been created by using Visual Studio 2017 and C# programming language. An image set of a patient is manually loaded to program. SQL server and MATLAB automatically takes the part to hold patient information written in DICOM format images in a database and detecting colon objects from whole CT colonography DICOM images. For the detection of colon objects on images, DICOM format images are read by the program and converted to grayscale images in Figure 8(a). Otsu thresholding method is applied to obtain binary images in Figure 8(b). To clear noise and fill small holes located in objects, imfill function is utilized in Figure 8(c). With the subtraction of the obtained image from binary image, colon objects with minor noisy objects and slab object stay at images in Figure 8(d). Average filter with 3x3 windows size is firstly applied on the obtained image. Median filter and Gaussian filter with 5x5 window size are separately applied. Connected component labeling is applied to give a label number for all objects on images. Areas of labeled objects with respect to pixels are calculated separately. Standard deviation of all calculated areas is calculated. Mean of areas which are less than standard deviation value is calculated. Then, objects whose areas are less than calculated mean value are removed from images. Therefore, minor noisy objects are removed from images. To remove slab object from images, K-means clustering method is used. Each object is classified into 2 clusters according to area size (K=2). Objects in a cluster which has greater mean of area values than another cluster is removed from images in Figure 8(e). Therefore, colon objects stay at image. Thanks to slice thickness information taken from DICOM images, all uploaded images of one patient are reconstructed in MATLAB environment. 3D imaging of colon is presented in CAS in Figure 8(f).

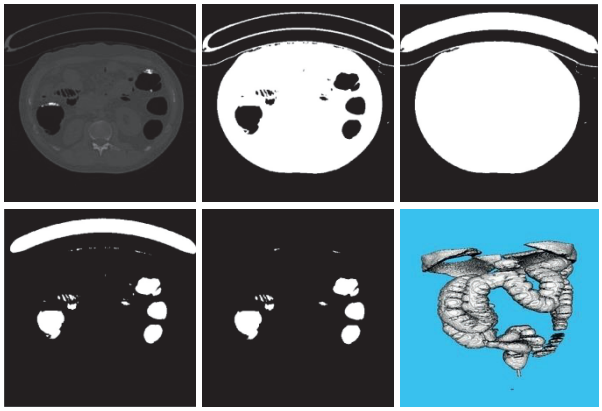


Figure 8: Processes applied on a sample image. (a) grayscale image, (b) binary image, (c) image obtained after filling holes, (d) image obtained after subtraction process, (e) image obtained after K-means clustering method, (f) 3D image.

After obtaining all 2D images and their reconstructed 3D volume image, parameters such as length of colon, area of colon, volume of colon, average color value of all colon pixels, entropy value of colon pixels calculated in Eq. (1), energy value of all colon pixels calculated in Eq. (2), contrast value of all colon pixels calculated in Eq. (3) and correlation value of all colon pixels calculated in Eq. (4) are presented as information in CAS [13].

$$Entropy = -\sum_{i=1}^{N}\sum_{j=1}^{N} p(i,j)\log p(i,j) \quad (1)$$

$$Energy = -\sum_{i=1}^{N}\sum_{j=1}^{N} p(i,j)^2 \quad (2)$$

$$Contrast = -\sum_{i=1}^{N}\sum_{j=1}^{N} (i-j)^2 p(i,j) \quad (3)$$

$$Correlation = -\sum_{i=1}^{N}\sum_{j=1}^{N} \left(\frac{i-\mu_x}{\sigma_x}\right)\left(\frac{j-\mu_y}{\sigma_y}\right) p(i,j) \quad (4)$$

Leap Motion controller is a motion tracking device that works by scanning hands and fingers with a high frame rate. Leap Motion controller can transfer position, motion and gesture information of hands to other software development environments via the Leap Motion API. The first element of Leap Motion API is the Frame object. This object contains hands, arms, fingers, gestures, and tools locating in frames detected by Leap Motion controller while scanning. They are also classes of frame object.

The hand's orientation is given by two vectors: first one is palm vector that points from the palm center towards the fingers, and second one is the palm normal that points out of the palm, perpendicular to the plane of the hand. In addition, thanks to the direction from the palm position to the fingers, the angle of pitching, yawing and rolling of the palm can be

calculated according to the horizontal plane (pitch, yaw, roll). On the other hand, the holding force is indicated by a parameter (grab) that varies between zero and one. If the hand is holding something or it is in the form of a fist the value is one otherwise the value is zero or close to zero. Palm direction vector and palm normal vector are given in Figure 9.

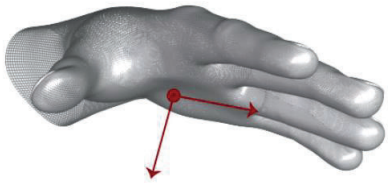


Figure 9: Palm normal and palm direction vectors [11].

Fingers are identified according to their types such as thumb, index finger, middle finger, ring finger and pinky. Leap Motion controller gives information about the type, bone structure, direction, tip position, length and width of each finger. Finger direction vectors are given in Figure 10.

Figure 10: Fingers and their direction vectors [11].

The Gesture class is a class expressing hand gestures. It has subclasses named as CircleGesture, SwipeGesture, KeyTapGesture and ScreenTapGesture. Thanks to Leap Motion API, hand movements are detected, processed by CAS, sent information to Arduino Nano and provided robotic arm with moving according human hand movements.

Robotic arm is designed to supply an easy control of servo motors. Thus, robotic arm parts are controlled via servo motors and Arduino nano that makes connection between robotic arm and CAS. There are 4 different servo motors providing robotic arm to move. First servo motor performs to move elbow part of robotic arm from up to down and vice versa. Second servo motor performs to move wrist part of robotic arm from up to down and vice versa. Third servo motor performs to move elbow part of robotic arm from left to right and vice versa. Fourth servo motor performs to open and close palm of robotic arm. The value read by Leap Motion controller and servo motor id are sent to Arduino Nano. According to the value and servo motors id, Arduino Nano controls and manages robotic arm with a correct movement.

Lattepanda single board computer is used as an embedded computer that connects all elements used in CAS. Lattepanda includes the developed windows form application and the database designed in SQL Server. Lattepanda is connected with Leap Motion controller, Arduino Nano and robotic arm. Instead of using a computer, Lattepanda is preferred because of transportation, taking a small place and being low-cost.

III. Results and Discussions

In this study, an automation system named as CAS is developed by concentrating on image processing to detect colon on CT colonography images in an automated manner, reconstruction models to view 3D colon, calculation of parameters related to colon, database system to keep information of patients and surgery progress and robotics arm hardware controlled by Leap Motion controller. After loading DICOM format images, application successfully detects colon objects with image processing techniques on 6524 images related to 7 patients. Each image contains 2D colon objects. CAS detects all of them. CAS reconstructs set of images based on slice thickness values and views 3D volume images of them. CAS calculates parameters related to colon objects. Tables designed in the database are held records for patients and their treatment history about colon. Touchless hardware control on 3D view of colon for surgical purpose and robotic arm performs accurately and successfully.

Recently, technology touches almost every sector to provide facilities, make lives easier, save time and so on. Robust and accurate automated systems in medicine is needed and significant for both patients and surgeons. CAS is a robust application developed to close the gap related to colon.

IV. Conclusion

CAS is a windows form application that automatedly detects colon in uploaded CT colonography DICOM images, views detected colons as a 3D volume image, calculates and presents parametric information to surgeon after processing images, enables surgeon to zoom, measure, rotate the viewing from different aspects without touching any computer parts thanks to Leap Motion controller for preoperative and intraoperative medical purpose. CAS is a helper tool for surgeons to view colon. By way of CAS, colon diseases can be examined. When CAS is improved by surgeons' requests and necessities, more convenient automation systems will be prepared for them. With such proper systems, healing percentage will increase. In this way, all sides of healthcare facilities like clinicians and patients will be satisfied with time saving diagnosis, situation analysis, surgeries with low infection risk and better healing rate.

Acknowledgment

Authors thank to cancer imaging archive to provide anonym CT colonography DICOM images.

References

[1] P B Riesz, "The life of Wilhelm Conrad Roentgen", *AJR Am J Roentgenol* vol 165 (6), pp 1533-1537, 1995

[2] W Röntgen, "On a new kind of rays", *Nature*, vol 53(1369), pp 274-276, 1896

[3] V Petrik, V Apok, J A Britton, B A Bell, M C Papadopoulos, "Godfrey Hounsfield and the Dawn of Computed Tomography", *Neurosurgery*, vol 58 (4), pp 780-787, 2006

[4] K Alemdaroğlu, T Akçal and D Buğra, *Kolon ve Rektum Anal Bölge Hastalıkları* İstanbul: Türk Kolon ve Rektum Cerrahisi Derneği, 2003

[5] Q K Al-Shayea and M S Al-Ani, "An efficient approach to 3d image reconstruction", *International Journal of Computer Science and Network Security*, vol 16(8), pp 35-41, 2016

[6] A Angelopoulou, A Psarrou, J G Rodríguez, S Orts, J Azorin-Lopez, K Revett, "3D reconstruction of medical images from slices automatically landmarked with growing neural models", *Neurocomputing*, vol 150, pp 16-25, 2015
[7] H Wang and X Pu, "3D medical CT images reconstruction based on VTK and Visual C++", *in Conf. Rec. 2009 3rd International Conference on Bioinformatics and Biomedical Engineering*, pp 1-4
[8] T Datino, A Arenal, M Pelliza, J Hernandez-Hernandez, F Atienza, E Gonzalez-Torrecilla, P Avila, L Bravo and F Fernandez-Avilez, "Comparison of the safety and feasibility of arrhythmia ablation using the amigo robotic remote catheter system versus manual ablation", *Am J Cardiol*, vol 113, pp 827-831, 2014
[9] D Stoianovici, "Robotic surgery", *World Journal of Urology*, vol 18, pp 289-295, 2000
[10] D B Camarillo, T M Krummel, J K Salisbury, "Robotic technology in surgery: past, present, and future", *The American Journal of Surgery*, vol 188(4), pp 2-15, 2004
[11] Leap Motion API 2020 Leap Motion Overview of API Available at: https://developer-archive leapmotion com/documentation/v2/csharp/devguide/Leap_Overview html [24 09 20]
[12] LattePanda 2020 Available at: https://www lattepanda com/products/3 html [24 09 20]
[13] R M Haralick, K Shanmugam and I Dinstein, "Textural Features for Image Classification", *IEEE Transactions on Systems, Man, and Cybernetics*, vol 3, pp 610–621, 1973

"Comparison of Deep Learning and Machine Learning Methods to Classify Retinal Images"

A. AKGUNDOGDU[1]

[1]Istanbul University - Cerrahpasa, Istanbul/Turkey, akgundog@istanbul.edu.tr

***Abstract* - Deep learning and machine learning are emerging as a powerful tool for analyzing medical images. Various deep learning and machine learning models are used to detect retinal disease using computer-aided diagnosis from retinal fundus images. In this study, retinal images of cataract patients are compared with normal retinal images and classified. A deep learning method and machine learning algorithms were used for classification in this study. As a result of this study, 91.5% success rate for DenseNet deep learning, and 79.5% success rate for the Random Forest method were obtained respectively. The success rates of the applied models were trained and validated to avoid overfitting by the cross-validation method.**

***Keywords* - Deep Learning, Machine Learning, Cataract, Retinal Abnormality**

I. Introduction

CATARACT is the cloudiness of the lens inside the eye that partially impairs vision and causes permanent vision loss if not detected and treated in the early stages. It is considered the most common cause of blindness. Cataracts present many challenges to patients' lives, such as appreciating changes in color and contracture, driving, reading, recognizing faces, and dealing with the glare from bright lights [1]. Early diagnosis of cataracts can reduce the risk of impaired vision and blindness. Cataract removal surgery is the most effective treatment and is usually performed in cases where cataracts negatively affect the patient's daily life and work. Although cataracts are not present in the retina, the clouding of the crystalline lens will reduce the light focused on the retina and cause deterioration of the fundus image quality. Experienced ophthalmologists can decide whether to undergo surgery by evaluating the difference between a non-cataract fundus image and a cataract image.

In this article, deep learning and machine learning methods are compared for cataract classification. 2D discrete wavelet method and histogram of directional gradients (HOG) are used in feature selection for machine learning. On the other hand, for deep learning, the densely connected neural network method was used. Finally, evaluation measures such as accuracy, sensitivity, specificity, precision, f-score, and area under the receiver operating characteristic curve (AUC) were utilized to measure the performance of used methods.

Investigation of fundus images for cataract detection has attracted the regard of numerous researchers over the past few years. [2-4]. The cataract detection system mainly consists of three stages: pre-processing, feature extraction, and classification. Various results have been noticing in the literature for detecting cataracts using retinal images. [4–8]. The automatic classification of retinal images for cataract detection has been studied in [2]. The use of CNN in computer-aided diagnosis for the detection and grading of cataracts on fundus images was performed in [9]. Higher-order empirical semantics and feature maps in the Pool5 layer were also examined.

CNN has been shown to perform better when the number of fundus images is high. However, collecting large numbers of real-time data sets poses a significant challenge. Also, pre-trained CNN is preferred to automatically extract features from medical images [10]. Zhang et al. utilized a deep convolutional neural network (DCNN) method for cataract classification in four classes [9]. A deeper network in combination with DCNN and random forests has been proposed by Ran et al. for the six-class cataract classification [11]. Xu et al. Presented the local-global feature representation using CNN and the deconvolution network [12].

The organization of the remaining article is as follows: Section II describes the dataset, feature extraction of retinal images, deep learning model, and machine learning classifiers. Experimental results are discussed in section III and the conclusion is displayed in section IV.

II. Material and Methods

The framework of machine learning and deep learning methods is given in Figure 1. The cataract classification framework consists of feature extraction, machine learning classifier, deep learning model, and performance evaluation. These steps are explained in the next sections.

A. Data set

The dataset used for the training and testing is collected from Kaggle database [13]. A total of 200 retinal images are labeled in two categories including normal (100) and cataract (100).

B. Machine learning with feature extraction

At this stage, to extract the feature vector from the images of retinal images; 2D discrete wavelet transform and HOG algorithm have been used.

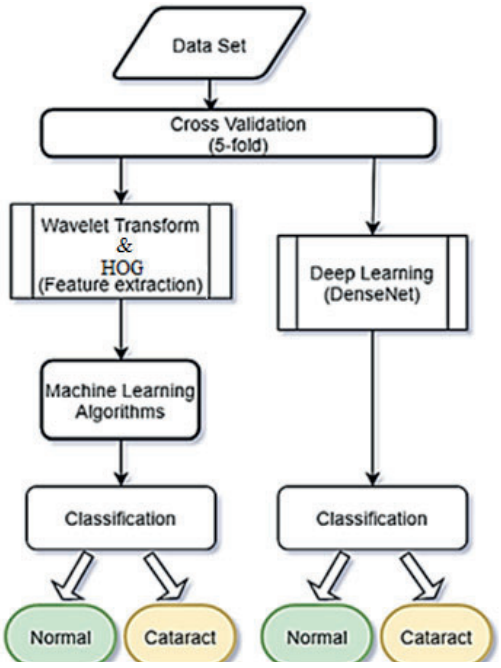


Figure 1: The framework of proposed models.

1. 2D Discrete Wavelet Transform

Wavelet transform is a preferred tool for obtaining the time-frequency representation of an image or a signal and provides successful results in many image analysis studies [14, 15]. Discrete wavelet transform separates the image into a lower resolution image, horizontal, vertical, and diagonal detail components. High-resolution subbands are found in edge and texture patterns in an image. This operation comes about in four decomposed subband images referred to low-low (LL), low-high (LH), high-low (HL), and high-high (HH). In this study, one-level 2D discrete Meyer wavelet transform type is used [16-18].

Wavelet decomposition of a signal helps to extract different information of the main signal such as high frequency or low frequency segments. This conversion is very useful for analyzing the signal or image at different scales or resolutions and provides highly successful results for classification. Wavelet transform has been used successfully in the classification of brain MRI and hyperspectral images in recent years [19,20].

Figure 2 shows a schematic chart of 2D discrete wavelet transform and vertical, horizontal, and diagonal states of cataract image.

The technique is very fast and can be used to extract basic structural information from an image. By applying wavelet transform, four subgroups are obtained as follows:

$$S^1 = S_a^1 + S_v^1 + S_h^1 + S_d^1 \quad (1)$$

The input image S^1 is decomposed into a first-level S_a^1 (approximation component, LL) and S_v^1, S_h^1, S_d^1 (detail components; vertical (LH), horizontal (HL), and diagonal (HH). Here, the retinal fundus image has been decomposed to 1- level using the DWT approach using Matlab wavelet analysis [21].

In the wavelet output block, 24 statistical properties of each image are subtracted as minimum (min), maximum (max), mean, standard deviation (1st moment, std), variance (var), and 3rd-moment values. These 24 features are the input of the machine learning model. The approximation, vertical, horizontal, and diagonals of the wavelet subbands with 24 features are shown in Figure 3.

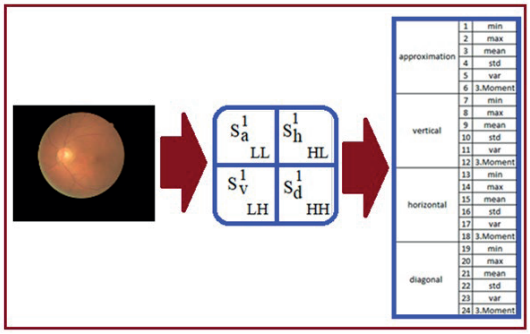


Figure 2: One-layer decomposition of 2D discrete wavelet transform and feature extraction block (24 statistical features of each image).

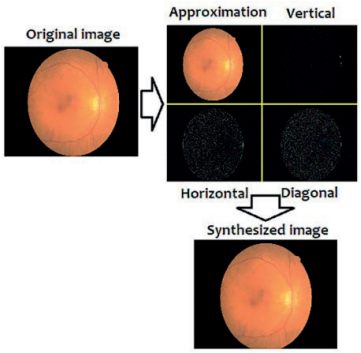


Figure 3: Cataract image and its 2D discrete wavelet coefficients at one-layer decomposition

2. Histogram of oriented gradients (HOG)

The use of this algorithm, which is a gradient-based feature extraction method, was first proposed by Shashua and Dalal [22]. This algorithm defines an image as a series of local histograms. Each of the local histograms is the distribution of the number of gradients that occur in specified directions, calculated from a specific area on the image defined as a cell.

The feature extraction from the images of the HOG method can be briefly summarized as follows: The image is parsed into small interconnected sections called cells. Larger structures are created that contain cell divisions called blocks. The orientation of the pixel values for each block is calculated. Inclination angles are obtained from these orientations. A gradient histogram is created from two different gradients obtained. The histogram obtained is subjected to a normalization process to minimize the slight differences in the image. The normalized histograms obtained for each block are added together to obtain

the feature vector of the image. The HOG method has been successfully implemented in face recognition in recent years [23,24].

C. Model Selection

In this work, Random Forest, Neural Network, Support Vector Machines, and k-Nearest Neighbor models, which are well-known machine learning methods in the literature, are used. These models are briefly summarized below.

1. Random Forest (RF)

It is the algorithm that increases the classification rate by generating more than one decision tree during the classification process [25]. Randomly selected decision trees come together to form the decision forest. Categorical with a large number of variables and class tags a distribution with variables, missing data, or unbalanced Works well for data sets that exhibit. Random decision forests correct the habit of decision trees to overfit training sets. Random forests generally outperform decision trees.

2. Support Vector Machines (SVM)

SVM is a supervised learning method with a learning algorithm that analyzes data used for classification and regression analysis problems [26, 27]. The basic SVM theory states that for a dataset that is sequentially separable without a perimeter of scores from two classes, the infinite number of lines dividing the classes is. The choice of a line that best separates the two classes is surrounded only by a subset of the training program known as support vectors. The boundary distance between support vectors is taken to represent the optimal decision. For problems where classes cannot be parsed as rotations, SVM uses an implicit transformation of the kernel function.

3. Neural Network (NN)

Neural network is computing processing systems that generally imitate the working principles of the human brain or central nervous system [28]. The first studies on this subject started with the modeling of the biological units that make up the brain and the application of neurons in computer systems. Neurons are connected by links, and each link has a weight value that expresses the importance of its input. These connections are called edges. Neurons and edges typically have a weight that changes as learning progresses. Weight increases or decreases the strength of the signal in a connect. Neurons can have a threshold at which a signal can only be sent if the total signal crosses this threshold. Typically, neurons gather in layers. Distinct layers can perform different transformations on their inputs. Signals travel from the first layer to the last layer, possibly after passing layers multiple times.

4. k- Nearest Neighbour (kNN)

kNN is a nonparametric method suggested by Thomas Cover and Hart used for classification and regression [30]. In both cases, the input comprises the closest training instances in the feature area. In the kNN classification, the output is a class membership. An object is classified by multiple votes of its neighbors, and object k is appointed to the class most joint among its closest neighbors (k is a positive integer). If $k = 1$, the object is simply assigned to that nearest neighbor's class. Euclidean or Mahalanobis distance is mostly used in distance measurements.

Figure 4 shows a block diagram of the different ML models compared in this study.

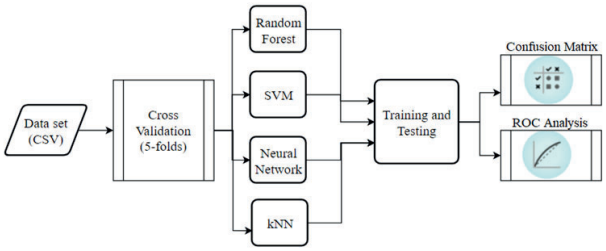


Figure 4: Knowledge flow of different ML models.

5. Deep Learning

When neural networks are trained, there is a decrease in feature maps due to convolution and subsampling processes. At the same time, there are losses in the image feature in the transition between layers. DenseNet (Densely connected neural network) system has been developed by Huang to use image feature information more effectively [31]. In the system, each layer is connected to the other layers in a feed-forward manner. In this way, any layer I can access the property information of all previous layers. It is possible to see the general network architecture of dense connected neural networks in Figure 5.

$[X_1, X_2, X_3, \ldots, X_{l-1}]$ layers are a combination of feature information. H_L is the transfer function used to process feature information. In this way, the rate of propagation of features to the network increases, re-use becomes easier and the number of parameters is greatly reduced. It is possible to see the densely connected neural network equation below.

$$X_l = H_l[X_0, X_1, X_2, \ldots, X_{l-1}] \quad (2)$$

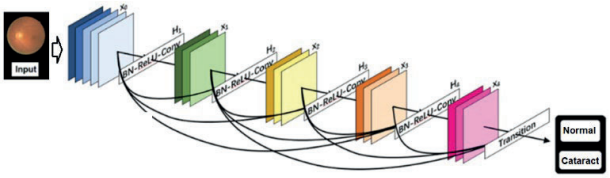


Figure 5: Densely connected neural network architecture [31].

D. Evaluation Criteria

The proposed methods are tested in a retinal dataset and performance criteria such as accuracy, sensitivity, and specificity, sensitivity is used to measure f-score and area under the curve (AUC) performance. Accuracy shows correctly classified samples. Sensitivity and specificity are correctly classified as true positive and true negative respectively. Precision indicates how many predicted are true positives. The f-score is the weighted average of recall (sensitivity) and precision. The mathematical formula of these metrics are as follows:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (3)$$

$$Sensitivity = \frac{TP}{TP+FN} \quad (4)$$

$$Specificity = \frac{TN}{TN+FP} \quad (5)$$

$$Precision = \frac{TP}{TP+FP} \quad (6)$$

$$f-Score = \frac{2*TP}{2*TP+FP+FN} \quad (7)$$

Samples are evaluated as true positive, true negative, false positive, and false negative according to the target value and predicted value. These are abbreviated respectively as TP, TN, FP, and FN.

E. Cross-Validation (CV)

In the cross-validation method, the *k* training set is divided into equal parts and one of them is used as the validation set. The remaining *k-1* pieces form the training set. A different section is used as the verification set each time. This process is repeated k times and each time its success in the validation set is calculated. Thus, *k* success values are obtained. The average *k* success values obtained are accepted as the success of the model.

III. Results and Discussion

In this study, 200 fundus image data sets consisting of cataract and normal images were used. These images are tagged for cataract (100) and normal (100) cases. 24 features are extracted from each image with the 2D-discrete wavelet method and 81 features are extracted from the HOG algorithm. These features are used for inputs to machine learning methods. At the output, two classes were determined as cataract and normal. The system designed for machine learning has a structure with 24 inputs and 2 outputs for wavelet transform and 81 inputs and 2 outputs for HOG algorithm. The confusion matrix results of the machine learning algorithms for 2D-discrete wavelet transform are given in Table 1, and the confusion matrix results of the machine learning algorithms for HOG are given in Table 2 respectively. Evaluation criteria for classification algorithms are obtained using 5-fold cross-validation. ROC curves of RF, SVM, NN, and kNN classifiers are seen in Figure 6 and Figure 7 respectively.

Table 1: The confusion matrix of the classification algorithms for 5-fold CV using 2D Discrete Wavelet.

Random Forest

Actual \ Predicted	cataract	normal	Σ
cataract	84	16	100
normal	25	75	100
Σ	109	91	200

Neural Network

Actual \ Predicted	cataract	normal	Σ
cataract	78	22	100
normal	27	73	100
Σ	105	95	200

SVM

Actual \ Predicted	cataract	normal	Σ
cataract	75	25	100
normal	28	72	100
Σ	103	97	200

kNN

Actual \ Predicted	cataract	normal	Σ
cataract	65	35	100
normal	25	75	100
Σ	90	110	200

Table 2: The confusion matrix of the classification algorithms for 5-fold CV using HOG.

Random Forest

Actual \ Predicted	cataract	normal	Σ
cataract	68	32	100
normal	24	76	100
Σ	92	108	200

Neural Network

Actual \ Predicted	cataract	normal	Σ
cataract	78	22	100
normal	22	78	100
Σ	100	100	20

SVM

Actual \ Predicted	cataract	normal	Σ
cataract	78	22	100
normal	20	80	100
Σ	98	102	200

kNN

Actual \ Predicted	cataract	normal	Σ
cataract	72	28	100
normal	26	74	100
Σ	98	102	200

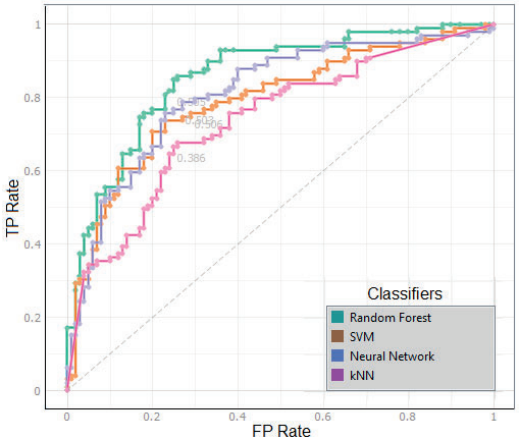


Figure 6: ROC curve of Random Forest, SVM, Neural Network, and kNN classifiers (5-fold CV using Wavelet).

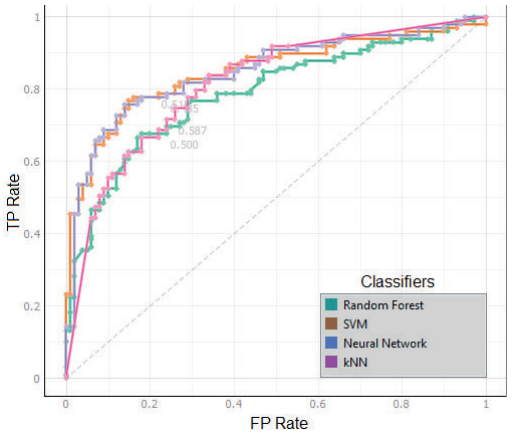


Figure 7: ROC curve of Random Forest, SVM, Neural Network, and kNN classifiers (5-fold CV using HOG).

The accuracy and loss curves for 5-fold cross-validation of the applied deep learning model are shown in Figure 8. The

confusion matrix obtained by using deep learning is seen in Table 3. The ROC curve of the deep learning classifier is shown in Figure 9.

Table 4 shows the values of the effective parameters used in machine learning models by using wavelet and HOG methods, and Table 5 shows the values of the parameters selected in the deep learning model. The classification results obtained from machine learning and deep learning methods compared with results are shown in Table 6. According to the results in Table 6, it is seen that the deep learning method gives the most successful results.

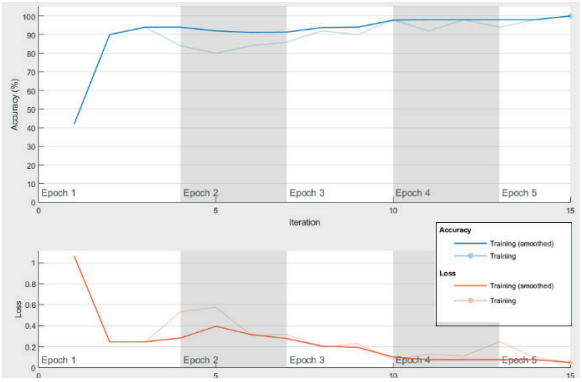


Figure 8: The accuracy and loss graph for 5-fold cross-validation of the deep learning model.

In this study, Matlab was used in feature extraction (Wavelet and HOG) and deep learning stages, and the Orange data mining tool was used in classification with machine learning.

Table 3: The confusion matrix of the classification algorithms for deep learning (5-fold CV).

Deep Learning (Densenet)

		Predicted		
		cataract	normal	Σ
Actual	cataract	90	10	100
	normal	7	93	100
	Σ	97	103	200

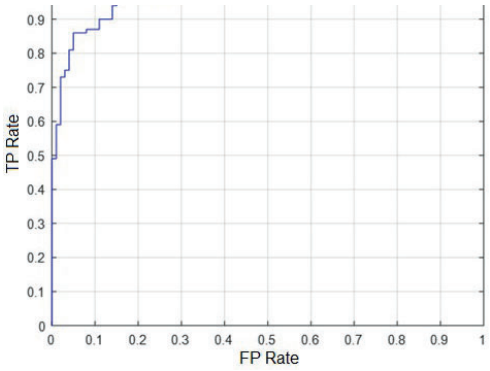


Figure 9: ROC curve of deep learning classifier.

Table 4: Defined parameters for machine learning models.

Classifier	Effective Parameters (HOG)	Effective Parameters (Wavelet)
Random Forest	Number of Trees: 18	Number of Trees: 20
Neural Network	Hidden Layer:3 Neurons: 10 16 10 Learning Rate 0 3 Activation Function: Relu	Hidden Layer:3 Neurons: 20 10 20 Learning Rate 0 3 Activation Function: Relu
SVM	C:2 Kernel: RBF	C:2 Kernel: RBF
kNN	Distance: Euclidean Number of Neighborhood k:3	Distance: Mahalanobis Number of Neighborhood k:3

Table 5: Defined parameters for the deep learning model.

Deep Learning	(DenseNet) Optimizer: Rmsprop Max epoch: 5 Iteration: 15 Stride: [2 2] Learning Rate: 0 001 Activation Functions: Relu & Softmax Number of layers: 201

Table 6: Comparison of classification performance measure indices using RF, NN, SVM, kNN, and deep learning on 5-fold CV.

Feature Extraction	Classifier	Sensitivity (%)	Specificity (%)	f-score	Accuracy (%)	ROC Area (AUC)
Wavelet	**Random Forest**	**0.84**	**0.75**	**0.804**	**0.795**	**0.859**
HOG		0 68	0 76	0 708	0 720	0 79
Wavelet	**Neural Network**	0 78	0 73	0 761	0 755	0 813
HOG		0 78	0 78	0 780	0 78	0 849
Wavelet	**SVM**	0 75	0 72	0 739	0 735	0 792
HOG		0 78	0 80	0 788	0 790	0 851
Wavelet	**kNN**	0 65	0 75	0 684	0 700	0 737
HOG		0 72	0 74	0 727	0 730	0 816
-	**Deep Learning**	**0.93**	**0.90**	**0.9163**	**0.915**	**0.956**

IV. Conclusion

In this study, 200 retinal fundus image data sets consisting of cataract and normal images were used. Machine learning and deep learning methods were compared according to the successful results obtained. In this study, RF, SVM, NN, and kNN algorithms were used for machine learning, and the DenseNet model was used for deep learning. The attributes for machine learning have been extracted from the 2D discrete wavelet transform and HOG algorithm. Experimental results show that the deep learning method outperforms machine

learning methods in terms of accuracy and other evaluation criteria. In this study, the time duration for feature extraction with machine learning is much shorter than deep learning, although deep learning yielded successful results.

REFERENCES

[1] D Allen, A Vasavada, "Cataract and surgery for cataract", *British Medical Journal* 333 (7559) 128, 2006

[2] U R Acharya, N Kannathal, E Ng, L C Min, J S Suri, "Computer-based classification of eye diseases", *IEEE International Conference of Engineering in Medicine and Biology Society*, pp 6121–6124, 2006

[3] M Yang, J -J Yang, Q Zhang, Y Niu, J Li, "Classification of retinal image for automatic cataract detection", in: IEEE International Conference on e-Health Networking, Applications & Services, pp 674–679, 2013

[4] J Zheng, L Guo, L Peng, J Li, J Yang, Q Liang, "Fundus image based cataract classification", *IEEE International Conference on Imaging Systems and Techniques*, pp 90–94, 2014

[5] L Guo, J -J Yang, L Peng, J Li, Q Liang, "A computer-aided healthcare system for cataract classification and grading based on fundus image analysis", *Comput. Ind* vol 69, 72–80, 2015

[6] J -J Yang, J Li, R Shen, Y Zeng, J He, J Bi, Y Li, Q Zhang, L Peng, Q Wang, "Exploiting ensemble learning for automatic cataract detection and grading", *Comput. Methods Prog. Biomed* vol 124 45–57, 2016

[7] L Xiong, H Li, L Xu, "An approach to evaluate blurriness in retinal images with vitreous opacity for cataract diagnosis", *J. Healthc.* Eng 2017

[8] Z Qiao, Q Zhang, Y Dong, J -J Yang, "Application of svm based on genetic algorithm in classification of cataract fundus images", *IEEE International Conference on Imaging Systems and Techniques* 1–5, 2017

[9] L Zhang, J Li, H Han, B Liu, J Yang, Q Wang, "Automatic cataract detection and grading using deep convolutional neural network", *IEEE International Conference on Networking, Sensing and Control* 60–65, 2017

[10] S S Aboutalib, A A Mohamed, M L Zuley, W A Berg, Y Luo, S Wu, "Do pre-trained deep learning models improve computer-aided classification of digital mammograms?" *Proc. SPIE* 10575, 2018

[11] J Ran, K Niu, Z He, H Zhang, H Song, "Cataract detection and grading based on combination of deep convolutional neural network and random forests" In: 2018 *International Conference on Network Infrastructure and Digital Content (IC-NIDC). IEEE*, Guiyang, China p 155–159

[12] X Xu, L Zhang, J Li, Y Guan, L Zhang, "A hybrid global-local representation CNN model for automatic cataract grading" *IEEE J Biomed Health Inform.* 24(2):556–567, 2019

[13] https://www kaggle com/jr2ngb/cataractdataset (accessed: 10 September 2020)

[14] S Chaplot, L M Patnaik, N R Jagannathan, "Classification magnetic resonance brain images using wavelets as input to support vector machine and neural network " *Biomedical Signal Processing and Control* 1:86-92, 2006

[15] D C R Pereira, P Ramos, M Z Nascimento, "Segmentation and detection of breast cancer in mammograms combining wavelet analysis and genetic algorithm " *Computer Methods and Programs in Biomedicine* vol: 114, pp 88–101, 2014

[16] Y Meyer, *Ondelettes et opérateurs: Ondelettes* Hermann ISBN 9782705661250, 1990

[17] Y Meyer, "Principe d'incertitude, bases hilbertiennes et algebres d'operateurs," *in: Séminaire Bourbaki*, vol 28, pp 209–223, 1985

[18] Y Meyer, "A review of an introduction to wavelets, by Charles K Chui, and Ten lectures on wavelets, by Ingrid Daubechies", *Bull. Amer. Math. Soc. (N.S.)* 28 350–360, 1993

[19] N Zikiou, M Lahdir & D Helbert, "Hyperspectral image classification using graph-based wavelet transform", *International Journal of Remote Sensing*, 41:7, 2624-2643, 2020

[20] Z Ullah, M U Farooq, S H Lee, D An, "A Hybrid Image Enhancement Based Brain MRI Images Classification Technique " *Med. Hypotheses*, 109922, 2020

[21] MathWorks "Matlab and Wavelet Toolbox 2013b " The MathWorks, Inc 2013

[22] N Dalal and B Triggs, "Histograms of oriented gradients for human detection " *In Computer Vision and Pattern Recognition*, 2005 CVPR 2005 IEEE Computer Society Conference on, volume 1, pages 886–893 IEEE, 2005

[23] J Xu, X Xue, Y Wu, et al , "Matching a composite sketch to a photographed face using fused HOG and deep feature models " *Vis Comput*, 2020

[24] M G Mohammed, A I Melhum, "Implementation of HOG Feature Extraction with Tuned Parameters for Human Face Detection" *International Journal of Machine Learning and Computing*, Vol 10, No 5, September 2020

[25] T K Ho, "Random Decision Forests " *Proceedings of the 3rd International Conference on Document Analysis and Recognition,* Montreal, QC, pp 278—282, 14–16 August 1995

[26] B E Boser, I M Guyon and V N Vapnik, "A Training Algorithm for Optimal Margin Classifiers," *Proc. Fifth Ann. Workshop Computational Learning Theory*, ACM Press, New York, pp 144-152, 1992

[27] V Vapnik, *The Nature of Statistical Learning Theory*, Springer-Verlag, New York, 1995

[28] J A Freeman and D M Skapura, *Neural Networks Algorithms , Applications and Programming Techniques.* New York, USA: Addison-Wesley Publishing Company, 1991

[29] M Negnevitsky, *Artificial Intelligence: A Guide to Intelligent Systems.* Pearson Education, 2005

[30] T Cover, P Hart, "Nearest neighbour pattern classification " *IEEE Trans. Inf. Theory* 13, 21–27, 1967

[31] G Huang, Z Liu, L V D Maaten, and K Q Weinberger, "Densely Connected Convolutional Networks " *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Hawaii-USA, 2261-2269, 21-26 July 2017

Classification of Chest Radiography Images Based on Deep Convolutional Neural Network for Coronavirus Disease Detection

C. KAYA[1]

[1] Department of Biomedical Engineering, Zonguldak Bulent Ecevit University, Zonguldak / Turkey,
ceren kaya@beun.edu.tr, crnkaya@hotmail.com

Abstract **- The new coronavirus disease (COVID-19), which started in China in December 2019, has spread rapidly among people living in other countries. According to daily statistics of the European Centre for Disease Prevention and Control, the disease approaches approximately 46,156,540 cases worldwide. Due to the increasing number of cases every day, a limited number of COVID-19 test kits are available in hospitals. Therefore, it is essential to implement an artificial intelligence-based detection system as a rapid alternative diagnosis option to prevent the spread of COVID-19 among humans. In this study, the performance comparison of two deep learning models used to efficiently diagnose coronavirus disease (COVID-19) on chest radiography (X-ray) images was performed. The data set used in the study consists of 219 COVID-19, 219 healthy and 219 viral pneumonia chest X-ray images of individuals. To classify these data, deep convolutional neural network with data augmentation and transfer learning models were used. The images in the data set were classified into three categories as COVID-19, viral pneumonia and healthy people using pre-trained ResNet50 and DenseNet201 deep learning models. The performances of these two deep learning models used in classification were compared in terms of accuracy and loss values. Performance results show that DenseNet201 pre-trained model provides higher accuracy for two different batch size values (94.7% accuracy for Batch Size = 2, 95.45% accuracy for Batch Size = 3) than ResNet50 model.**

Keywords **- Coronavirus Disease, Viral Pneumonia, Chest X-ray Radiographs, Convolutional Neural Network, Deep Transfer Learning, Data Augmentation.**

I. INTRODUCTION

THE coronavirus disease (COVID-19) outbreak occurred in Wuhan, China, in December 2019. It has since become a serious public health problem worldwide [1, 2]. The number of deaths from pneumonia caused by the COVID-19 virus is increasing day by day. One of the most important methods used in the diagnosis of pneumonia in the world is chest radiography (X-ray) that is a fast, inexpensive cost [3, 4] and widely used clinical method [5-10]. In chest X-ray method, the patient is given a lower radiation dose compared to computed tomography (CT) and magnetic resonance imaging (MRI) [11].

In literature, studies about the diagnosis of COVID-19 disease by using chest X-rays have binary or multiple classifications applied with different dataset. Raw data is used in some of the studies while others have feature extraction process. Among the studies, the most preferred method is deep convolutional neural network (CNN). Khan et al. classified chest X-ray images from normal, bacterial, and viral pneumonia cases using Xception model for the detection of COVID-19 infection [12]. Ghoshal and Tucker used dropweights based Bayesian CNN model using chest X-ray images for the diagnosis of COVID-19 [13]. Hemdan et al. built VGG19 and DenseNet models to diagnose COVID-19 from X-ray images [14]. Ucar and Korkmaz worked on X-ray images for COVID-19 diagnosis and implemented SqueezeNet model with Bayesian optimization [15].

In this study, we have proposed an automatic prediction of COVID-19 using a deep convolutional neural network based ResNet50 and DenseNet201 pre-trained transfer models to obtain higher prediction accuracies for the dataset containing chest X-ray images of normal (healthy), COVID-19, and viral pneumonia patients.

II. CHEST X-RAY DATASET

In this study, chest X-ray images of 219 normal (healthy), 219 viral pneumonia and 219 COVID-19 patients have been obtained from the open source Kaggle repository called "COVID-19 Radiography Database" [16]. Distribution of images per class in used dataset is given in Table 1.

Table 1: Number of chest X-ray images per class.

COVID-19		Viral Pneumonia
219	219	219

All images were resized to 224x224 pixel size. In Figure 1, representative chest X-ray images of normal (healthy), COVID-19 and viral pneumonia patients are given, respectively.

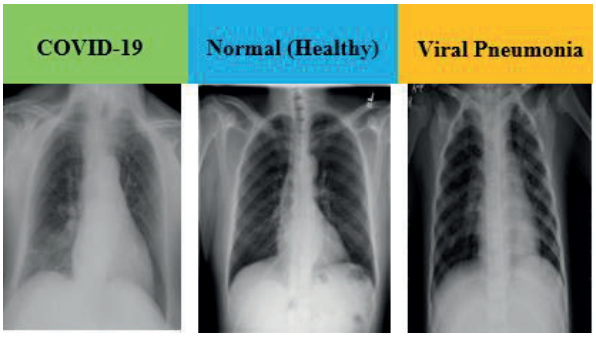


Figure 1: Representative chest X-ray image of COVID-19 (first column), normal (healthy) (second column) and viral pneumonia (third column) patient.

III. PRE-TRAINED DEEP LEARNING MODELS

One of the biggest drawbacks researchers face in analyzing medical data is the limited number of data sets. While deep learning models usually require a lot of data, the biggest advantage of using the deep transfer learning method is that it allows training of data with fewer datasets and requires less computational costs. With the transfer learning method, which is widely used in the deep learning field, the information gained from the pre-trained model on a large data set is transferred to the model to be trained [17]. In this study, we built deep convolutional neural network (CNN) based ResNet50 and DenseNet201 models for the classification of COVID-19 chest X-ray images. In addition, we applied transfer learning technique that was realized by using ImageNet data to overcome insufficient data and training time. ResNet50 [18] and DenseNet201 [19] deep learning models were 50-layer and 201-layer networks, respectively that were trained on the ImageNet dataset. ImageNet is an image database with more than 14 million images belonging to more than 20 thousand categories created for image recognition competitions [20].

IV. EXPERIMENTAL RESULTS

All experiments were performed on Python programming-based Google Colaboratory (Colab) Linux server having Ubuntu 18.04 operating system using Tesla K80 Graphics Processing Unit (GPU). ResNet50 and DenseNet201 CNN models were pre-trained with random initialization weights by optimizing cross-entropy function with adaptive moment estimation (ADAM) optimizer ($\beta1 = 0.9$ and $\beta2 = 0.999$). The learning rate and number of epochs were experimentally set to 1e-5 and 40, respectively for all experiments. In addition, classification performances of the models were tested separately by setting batch size 2 and 3. The dataset used was randomly split into two independent datasets with 80% and 20% for training and test, respectively.

A. ResNet50 Model (Batch Size = 2)

Training / test accuracy and loss values for Batch Size = 2 of pre-trained ResNet50 model are given in Figure 2 and Figure 3, respectively.

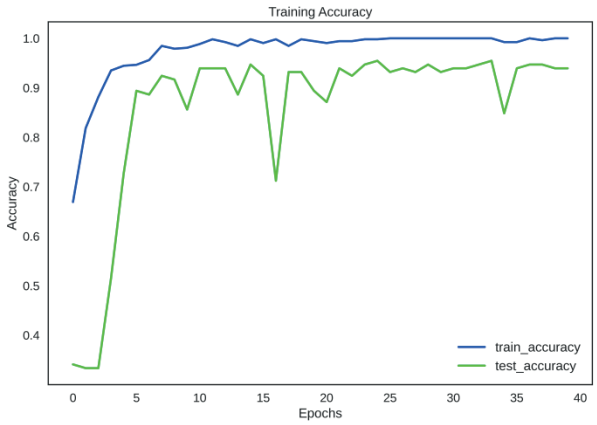


Figure 2: Training and test accuracy values of ResNet50 model for Batch Size = 2.

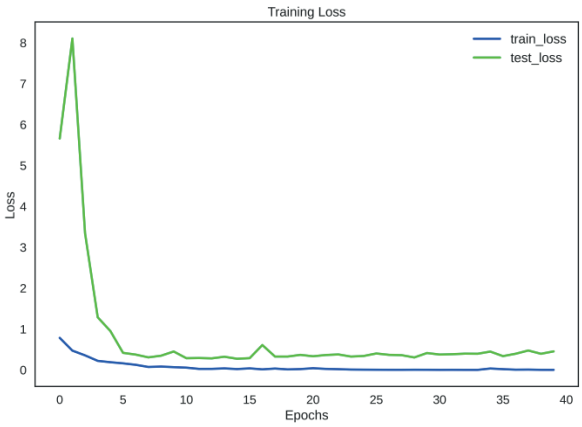


Figure 3: Training and test loss values of ResNet50 model for Batch Size = 2.

In Figure 4, confusion matrix of COVID-19, normal and viral pneumonia test results obtained from ResNet50 model for Batch Size = 2 is given. Besides the confusion matrix, receiver operating characteristic (ROC) curve plots and areas for each class obtained from ResNet50 model for Batch Size = 2 are given in Figure 5.

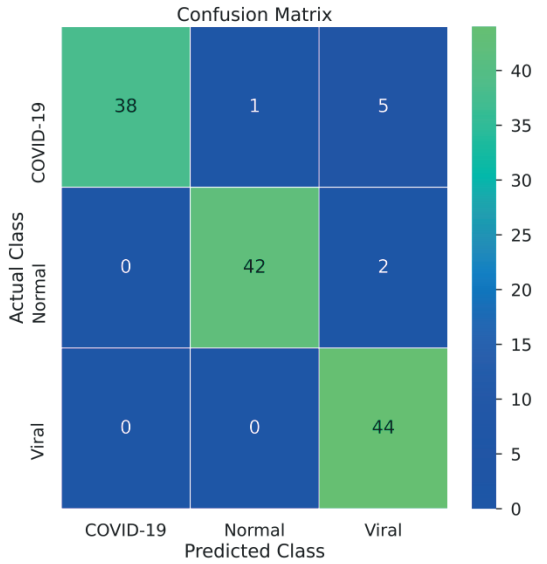


Figure 4: Confusion matrix obtained from ResNet50 model for Batch Size = 2.

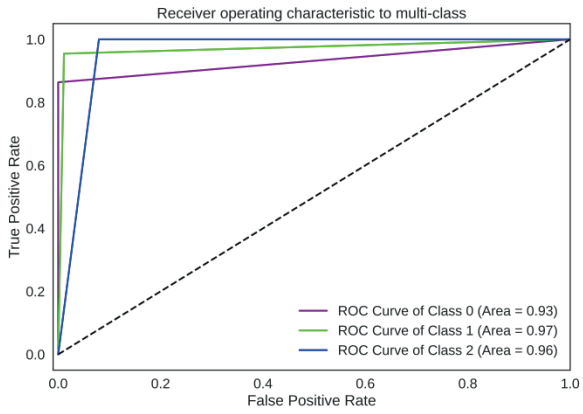


Figure 5: ROC curve plots obtained from ResNet50 model for Batch Size = 2.

B. DenseNet201 Model (Batch Size = 2)

Training / test accuracy and loss values for Batch Size = 2 of pre-trained DenseNet201 model are given in Figure 6 and Figure 7, respectively.

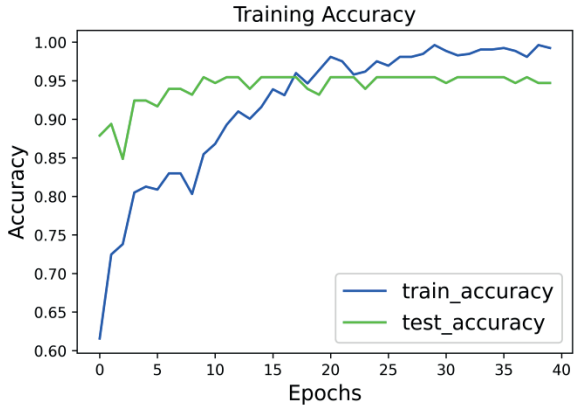


Figure 6: Training and test accuracy values of DenseNet201 model for Batch Size = 2.

Figure 7: Training and test loss values of DenseNet201 model for Batch Size = 2.

In Figure 8, confusion matrix of COVID-19, normal and viral pneumonia test results obtained from DenseNet201 model for Batch Size = 2 is given. Besides the confusion matrix, receiver operating characteristic (ROC) curve plots and areas for each class obtained from DenseNet201 model for Batch Size = 2 are given in Figure 9.

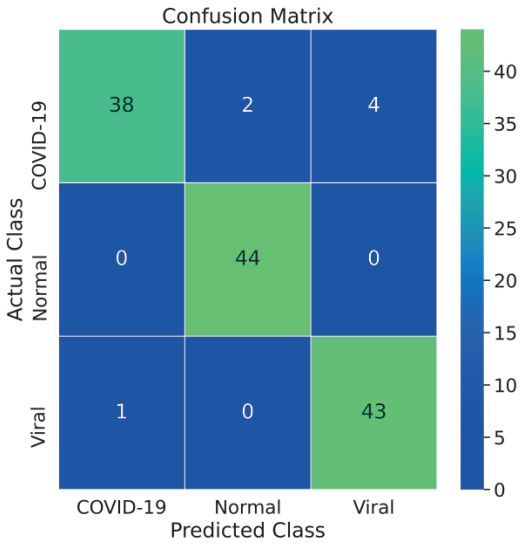


Figure 8: Confusion matrix obtained from DenseNet201 model for Batch Size = 2.

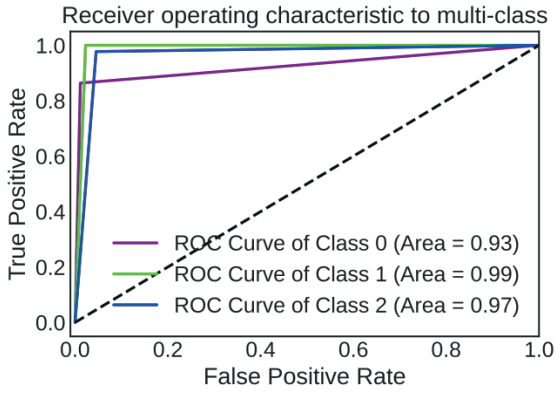


Figure 9: ROC curve plots obtained from DenseNet201 model for Batch Size = 2.

C. *ResNet50 Model (Batch Size = 3)*

Training / test accuracy and loss values for Batch Size = 3 of pre-trained ResNet50 model are given in Figure 10 and Figure 11, respectively.

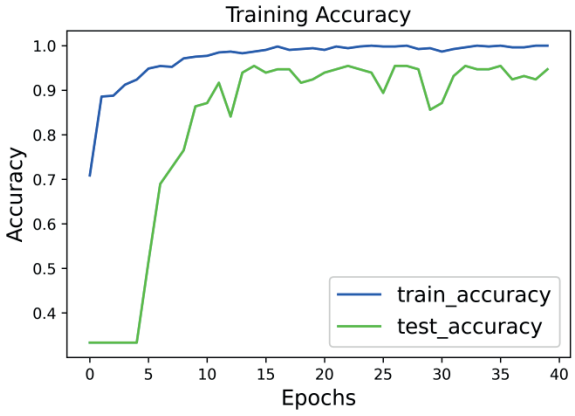


Figure 10: Training and test accuracy values of ResNet50 model for Batch Size = 3.

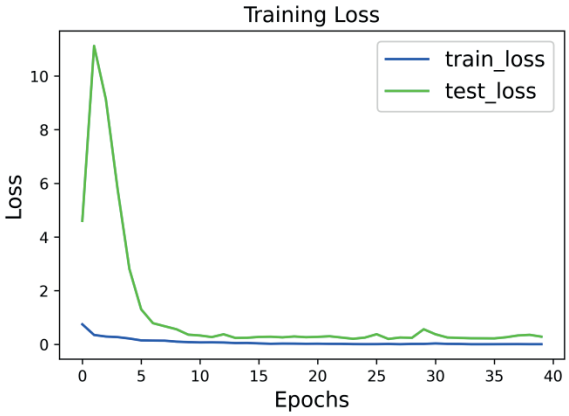


Figure 11: Training and test loss values of ResNet50 model for Batch Size = 3.

In Figure 12, confusion matrix of COVID-19, normal and viral pneumonia test results obtained from ResNet50 model for Batch Size = 3 is given. Besides the confusion matrix, receiver operating characteristic (ROC) curve plots and areas for each class obtained from ResNet50 model for Batch Size = 3 are given in Figure 12.

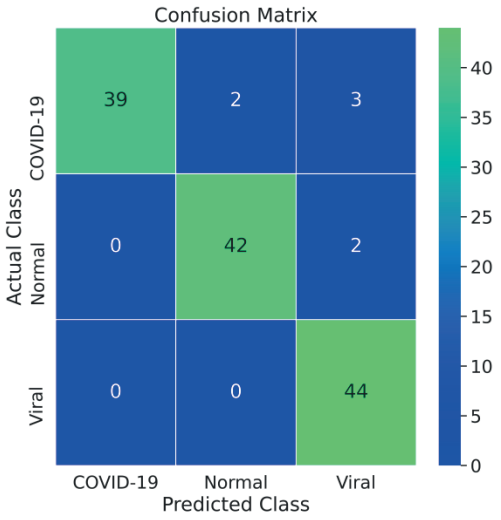


Figure 12: Confusion matrix obtained from ResNet50 model for Batch Size = 3.

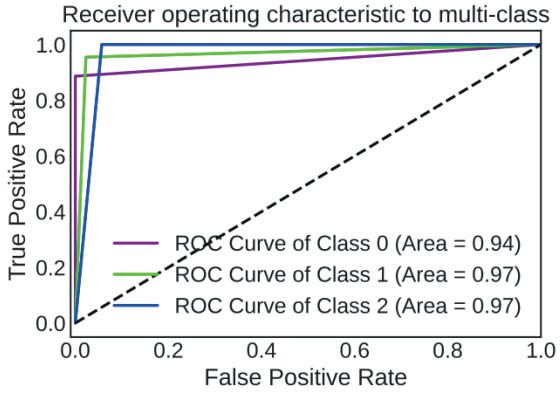


Figure 13: ROC curve plots obtained from ResNet50 model for Batch Size = 3.

D. DenseNet201 Model (Batch Size = 3)

Training / test accuracy and loss values for Batch Size = 3 of pre-trained DenseNet201 model are given in Figure 14 and Figure 15, respectively.

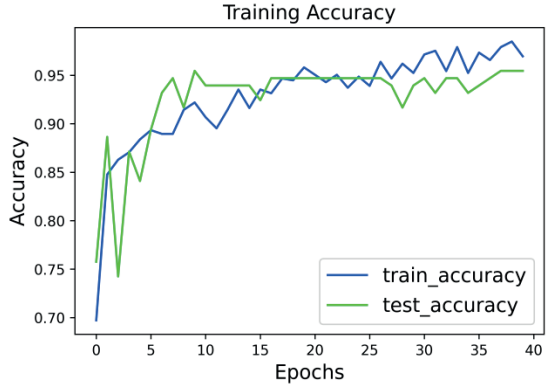


Figure 14: Training and test accuracy values of DenseNet201 model for Batch Size = 3.

Figure 15: Training and test loss values of DenseNet201 model for Batch Size = 3.

In Figure 16, confusion matrix of COVID-19, normal and viral pneumonia test results obtained from DenseNet201 model for Batch Size = 3 is given. Besides the confusion matrix, receiver operating characteristic (ROC) curve plots and areas for each class obtained from DenseNet201 model for Batch Size = 3 are given in Figure 17.

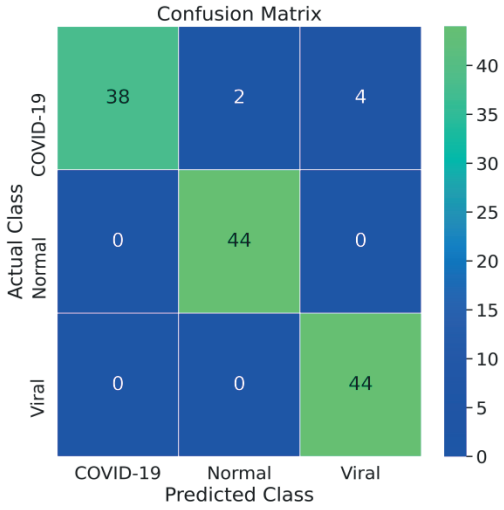


Figure 16: Confusion matrix obtained from DenseNet201 model for Batch Size = 3.

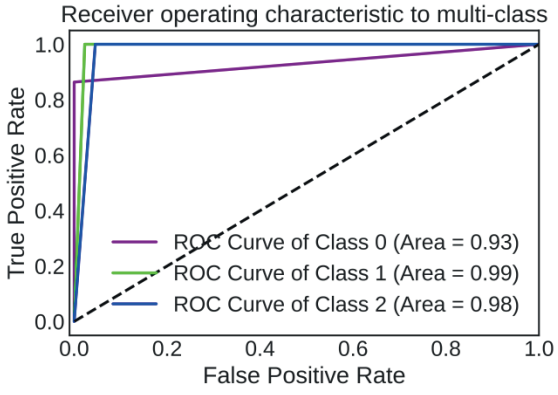


Figure 17: ROC curve plots obtained from DenseNet201 model for Batch Size = 3.

Overall accuracy evaluation metric was used for the calculation of classification performances of used deep transfer learning models. In another detailed performance, comparisons of two models using the test data are shown in Table 2. We have obtained better performance as overall accuracy of 94.7% (Batch Size = 2) and 95.45% (Batch Size = 3) for DenseNet201 pre-trained model than overall accuracy of 93.94% (Batch Size = 2) and 94.7% (Batch Size = 3) for ResNet50 pre-trained model. DenseNet201 model provides superiority over ResNet50 model for each batch size, training, and testing stage.

Table 2: Overall classification accuracy performance results obtained from two pre-trained CNN models.

Models / Batch Size	ResNet50 (Batch Size = 2)	DenseNet 201 (Batch Size = 2)	ResNet50 (Batch Size = 3)	DenseNet 201 (Batch Size = 3)
Overall Accuracy	93.94%	94.7%	94.7%	95.45%

V. CONCLUSION

Early prediction of COVID-19 disease is crucial to prevent spread among people. In this study, we proposed a deep transfer learning-based approach that uses chest X-ray images from normal, COVID-19 and viral pneumonia patients to automatically predict COVID-19 patients. In the comparison of two different deep learning models with two different batch size values, 525 chest X-ray images in the data set were used to train the deep learning models and the remaining 132 images were used to test the models. According to the performance results, pre-trained DenseNet201 model provides higher classification performance (94.7% accuracy for Batch Size=2, 95.45% accuracy for Batch Size=3) than pre-trained ResNet50 model (93.94% accuracy for Batch Size=2, 94.7% accuracy for Batch Size=3). According to the study findings, it is believed that higher classification performance obtained from the models will help radiologists make decisions in clinical practice. This study provides insight into how deep transfer learning methods can be used to detect COVID-19 at an early stage. In future studies, the classification performance of different CNN models will be tested by increasing the number of COVID-19 chest X-ray images in the data set.

REFERENCES

[1] K Roosa, Y Lee, R Luo, A Kirpich, R Rothenberg, J M Hyman, and et al , "Real-time forecasts of the COVID-19 epidemic in China from February 5th to February 24th, 2020", *Infect. Dis. Model* , vol 5, pp 256–263, 2020

[2] L Yan, H-T Zhang, Y Xiao, M Wang, C Sun, J Liang, and et al , "Prediction of criticality in patients with severe COVID-19 infection using three clinical features: a machine learning-based prognostic model with clinical data in Wuhan", *medRxiv* 2020 02 27 20028027, 2020

[3] A K Jaiswal, P Tiwari, S Kumar, D Gupta, A Khanna, and J J Rodrigues, "Identifying pneumonia in chest X-rays: A deep learning approach", *Measurement*, vol 145, pp 511–518, 2019

[4] B Antin, J Kravitz, and E Martayan (2017) Detecting Pneumonia in Chest X-Rays with Supervised Learning [Online] pp 1–5 Available: http://cs229 stanford edu/proj2017/final-reports/5231221 pdf

[5] N N Das, N Kumar, M Kaur, V Kumar, and D Singh, "Automated Deep Transfer Learning-Based Approach for Detection of COVID-19 Infection in Chest X-rays", *IRBM*, 2020, Available: https://doi org/10 1016/j irbm 2020 07 001

[6] E Ayan, and H M Ünver, "Diagnosis of Pneumonia from Chest X-Ray Images Using Deep Learning", in Conf Rec 2019 Scientific Meeting on Electrical-Electronics Biomedical Engineering and Computer Science (EBBT), pp 1-5, Available:https://doi org/10 1109/EBBT 2019 8741582

[7] P Afshar, S Heidarian, F Naderkhani, A Oikonomou, K N Plataniotis, and A Mohammadi, "COVID-CAPS: A Capsule Network-based Framework for Identification of COVID-19 cases from X-ray Images", *arXiv*:2004 02696v2, 2020

[8] N S Punn, and S Agarwal, "Automated diagnosis of COVID-19 with limited posteroanterior chest X-ray images using fine-tuned deep neural networks", *arXiv*:2004 11676v2, 2020

[9] M Barstugan, U Ozkaya, and S Ozturk, "Coronavirus (COVID-19) Classification using CT Images by Machine Learning Methods", *arXiv*:2003 09424, 2020

[10] A Narin, C Kaya, and Z Pamuk, "Automatic Detection of Coronavirus Disease (COVID-19) Using X-ray Images and Deep Convolutional Neural Networks", *arXiv*:2003 10849, 2020

[11] G Gaál, B Maga, and A Lukács, "Attention U-Net Based Adversarial Architectures for Chest X-ray Lung Segmentation", *arXiv*:2003 10304, 2020

[12] A I Khan, J L Shah, and M M Bhat, "Coronet: A deep neural network for detection and diagnosis of COVID-19 from chest x-ray images", Comput Meth Prog Bio, vol 196, pp 105581, 2020

[13] B Ghoshal, and A Tucker A, "Estimating Uncertainty and Interpretability in Deep Learning for Coronavirus (COVID-19) Detection", *arXiv*:2003 10769v1, 2020

[14] E E D Hemdan, M A Shouman, and M E Karar, "COVIDX-Net: A Framework of Deep Learning Classifiers to Diagnose COVID-19 in X-Ray Images", *arXiv*:2003 11055, 2020

[15] F Ucar, and D Korkmaz, "COVIDiagnosis-Net: Deep Bayes-SqueezeNet based diagnosis of the coronavirus disease 2019 (COVID-19) from X-ray images", *Med. Hypotheses*, vol 140, pp 109761, 2020

[16] T Rahman COVID-19 Radiography Database (2020) Available: https://www kaggle com/tawsifurrahman/covid19-radiography-database

[17] N Jmour, S Zayen, and A Abdelkrim, "Convolutional neural networks for image classification", in *Conf. Rec. 2018 Int. Conf* on Advanced Systems and Electric Technologies (IC_ASET), pp 397–402

[18] Z Wu, C Shen, and A Van Den Hengel, "Wider or Deeper: Revisiting the ResNet Model for Visual Recognition", *Pattern Recognit.*, vol 90, pp 119–133, 2019

[19] K Sahinbas, and F O Catak (2020) Transfer Learning Based Convolutional Neural Network for COVID-19 Detection with X-Ray Images Available: https://www ozgurcatak org/files/papers/covid19-deep-learning pdf

[20] O Russakovsky, J Deng, H Su, J Krause, S Satheesh, S Ma S, and et al , "ImageNet Large Scale Visual Recognition Challenge", *Int. J. Comput. Vis.*, vol 115, pp 211–252, 2015

Quantification of Finger Tapping Variability by an Embedded System Implementing Poincare Analysis in Real-Time

S. AKSOY[1] and G. ERTAS[2]

[1] Graduate School of Natural and Applied Sciences, Yeditepe University, Istanbul/Turkey
[2] Department of Biomedical Engineering, Yeditepe University, Istanbul/Turkey, gokhan.ertas@yeditepe.edu.tr

***Abstract* – Development of an embedded system implementing Poincare analysis in real-time to quantify finger tapping variability is aimed. Finger-tapping signals demonstrating large, small, and very small response initiation times are generated synthetically. The embedded system is designed using an Arduino Micro microcontroller unit determining the response initiation time estimates from the signals and calculating the Poincare parameters SD1, SD2, and SD1/SD2 for the estimates. The signals, estimates, and parameters are also analyzed on a personal computer for validation. The embedded system outputs the same results as obtained on the computer. As the variation in response initiation time increases, SD1 and SD2 increase systematically, however, no systematic change is observed for SD1/SD2. Real-time calculation of Poincare parameters by embedded systems is feasible and offers simultaneous quantification of finger-tapping variability that opens up new opportunities for therapeutic devices. Additional work is needed to figure out the potential benefits in practice.**

***Keywords* - Finger-tapping, quantification, Poincare analysis, real-time, embedded system.**

I. Introduction

FINGER tapping is influenced by visual and sound-related incentives, emotional and physical health, and other elements affecting nervous and skeletal systems [1]. Since its introduction in 1947 by Ward Halstead in 1947, the test has been preferably used in evaluating the muscle control and motor ability in the upper extremities for detecting neurological disturbances such as in Parkinson's disease, and Alzheimer's malady [2]. The test has been also used in neuropsychiatry by facilitating the prediction of reaction time, cognitive impairment, and evaluation of hand skills and coordinates [3]. Evaluation of specific characteristics or changes in the finger-tapping pattern over time is often performed qualitatively by visual observations that can be advanced when an additional instrument such as a mechanical counter or keyboard is adapted during the finger-tapping test. However, the success of the evaluation relies highly on the observer experience and is quite error-prone accordingly when compared to quantitative evaluation.

Quantitative evaluation of finger-tapping patterns involves the detection of finger tapping pattern by a sensing setup and extraction of temporal measures of tapping responses from the tapping pattern captured by the setup. The sensor can be a force-sensitive resistor, an accelerometer, or a magnetometer [4-6] and the measures may be the time from the onset of one finger tap until the offset of the same finger tap (response duration time), the time from the offset of one finger tap until the onset of the next finger tap (response initiation time), and the total number of finger taps per unit of time (response rate) [7]. The mean and standard deviation values of these measures have been studied commonly. For instance, the response rate is lower in the elderly adults compared with young adults on average [8]. However, limited information on finger tapping is obtained when the average or standard deviation values are just considered.

For advanced evaluations from finger-tapping, the use of sophisticated methods involving analyses of the tapping signal in time, frequency, or wavelet domains are gaining increasing interest. These methods offer detailed spectral analyses of the finger-tapping patterns [9]. In this study, we aim to develop an embedded system implementing Poincare analysis in real-time to quantify finger-tapping variability.

II. Materials and Methods

The work is carried out by simulations performed using the tools provided by Proteus Design Suite (Labcenter Electronics Ltd, England). An embedded system equipped with an Arduino Micro microcontroller unit is designed and programmed to apply Poincare analysis on finger-tapping signals in real-time. The Easy HDL generator tool is used to generate the finger-tapping signals, the virtual oscilloscope tool is used to visualize the signals generated, and the virtual terminal tool is used to output the results of Poincare analysis. Figure 1 shows the electronic circuit connections between the microcontroller and the tools mentioned.

A. Generation of Finger-Tapping Signal

The behavior of a one-touch style switch relying on capacitive sensing technology in capturing finger-tapping signal is considered [10]. The switch is thought to be equipped with circuitry to minimize the bouncing artifact and to output a digital rectangular pulse as a response to touch. While a finger touching event is occurring, the pulse amplitude is logic high (+5Vdc) and it is logic zero (0V) for no-touch.

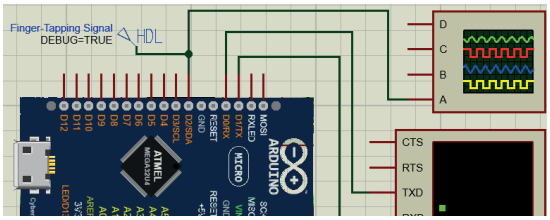


Figure 1: Electronic circuit connections.

By analyzing the finger-tapping signal generated, one may estimate response duration time, response initiation time, and the response rate measures for a finger-tapping sequence. In the current study, demonstrative signals are generated synthetically to mimic varying levels for the response initiation time but constant response duration time and response rate. For the purpose, a script is written for the Easy HDL generator offered by the simulation environment. The script expresses a finger tapping signal with a period of 550ms. The finger touch is mimicked with a rectangular pulse with a constant duration time of 50ms and amplitude of 5V. The pulse is superimposed on the signal for a period considering a specific time offset, t_{offset} (One-period signal obtained for t_{offset}= 250ms is as illustrated in Figure 2). Different levels of randomness are implemented for t_{offset} to obtain three different signals illustrating response initiation time with large, small, and very small variations.

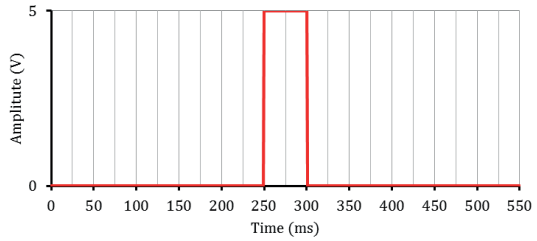


Figure 2: One-period finger tapping signal with t_{offset}= 250ms.

B. Estimation of the Response Initiation Time

A microcontroller program is developed using Arduino IDE to capture and process the finger-tapping signal fed to the digital I/O pin, D2 of the microcontroller unit. The program detects the transits from logic 0 to logic 1 on the signal and estimates the time elapsed between two consecutive transits detected (i.e. the response initiation time). The microcontroller stores the estimates and following the twenty-fifth estimation; it performs Poincare analysis for the estimates stored.

C. Calculation of the Poincare Parameters

The microcontroller implements Poincare analysis using the ellipse fitting technique and calculates the Poincare parameters SD1 and SD2, and the ratio of SD1 to SD2. Let N be the total number of response initiation time estimates and let RR_n denote the n-th response initiation time estimate ($n = 1,2, \dots, N-1$). $SD1$ and $SD2$ are calculated using [11]:

$$SD1 = \frac{1}{\sqrt{2}} std(RR_n - RR_{n+1}) \quad (1a)$$

$$SD2 = \sqrt{2 * std(RR_n)^2 - SD1^2} \quad (1b)$$

The microcontroller transfers the Poincare parameters calculated via its serial port to the virtual terminal offered by the simulation environment to inform the user.

D. Validation

The finger-tapping signal generated, the response initiation times estimated and the Poincare parameters calculated by the microcontroller are validated using Matlab (v8.2; Natick, MA). All simulations and validations are performed on a standard desktop PC (Intel i5-4460 3.20GHz processor, 12GB memory, and 64-bit operating system).

III. Results

Three representative finger-tapping signals generated to mimic response initiation time with large, small, and very small variations are as seen in Figure 5. Descriptive statistics

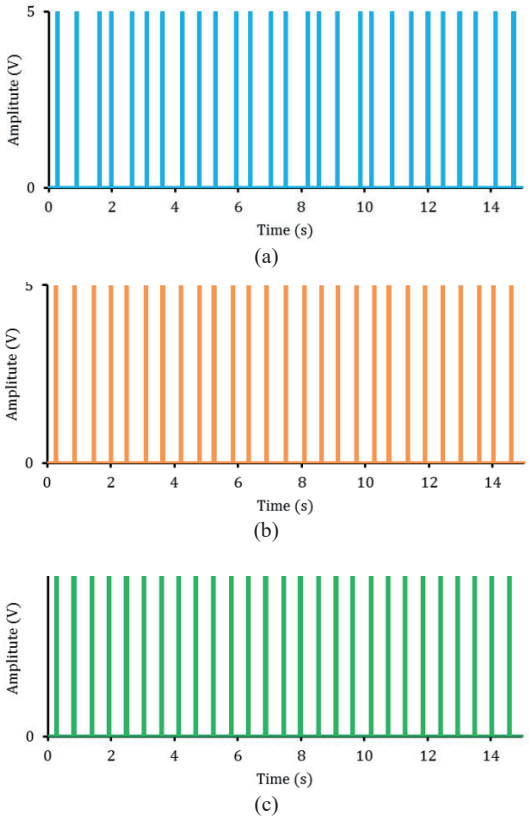


Figure 5: Finger-tapping signal possessing response initiation time with (a) large, (b) small, and (c) very small variations.

for the response initiation time are 305-660ms range and 501.3ms average, and 80.4ms standard deviation for the signal possessing large variations, 391-577ms range, 501.3ms average, and 43.9ms standard deviation for the signal exhibiting small variations, and 468-524ms range, 500.3ms average and 16.1ms standard deviation for the signal unveiling very small variations (A lower standard deviation indicates a smaller variation).

Figure 6 shows the Poincare plots for the response initiation time estimates from the signals obtained by the microcontroller. Poincare parameters calculated by the microcontroller for the estimates are as listed in Table 1. As the variation increases, *SD*1 and *SD*2 increase, however, no increase or decrease is observable for *SD*1/*SD*2. The estimates and the Poincare parameter values by the microcontroller are identical to the ones computed by using the personal computer.

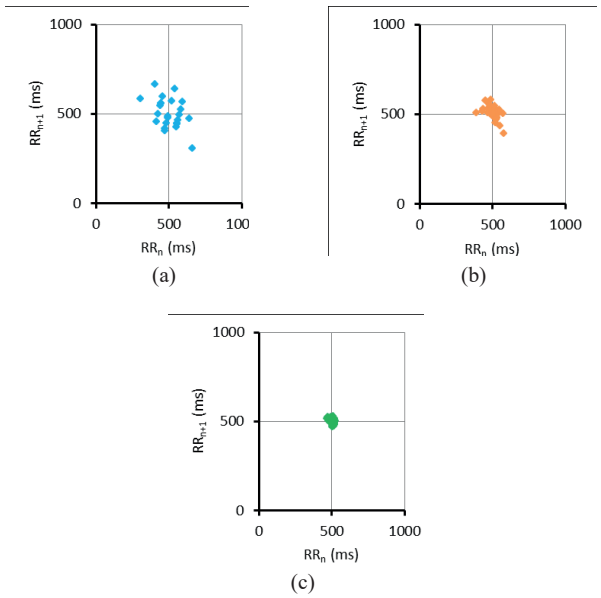


Figure 6: Poincare plots for the response initiation time with (a) large variation, (b) small variation, and (c) very small variation.

Table 1: Poincare parameter values obtained.

Finger-tapping signal exhibiting	*SD*1	*SD*2	*SD*1/*SD*2
Large variation	97.8	58.0	1.7
Small variation	53.9	30.8	1.8
Very small variation	19.4	11.9	1.6

Ten different finger-tapping signals are generated for each variation degree for the response initiation time randomly. Table 2 lists the average values for the Poincare parameters calculated by the microcontroller unit for these signals. As expected, there are significant differences for different levels of variation between the *SD*1 and *SD*2 parameters however no significant differences are present between the *SD*1/*SD*2 parameter. As the variation increases, *SD*1 and *SD*2 increase, however, no increase or decrease is present for *SD*1/*SD*2.

Table 2: Average values for the Poincare parameters.

Finger-tapping s	*SD*1	*SD*2	*SD*1/*SD*2
Large variation	98.5±8.6 [a]	67.5±6.5	1.5±0.2
Small variation	60.6±5.3	35.2±5.4	1.8±0.4
Very small variation	18.5±2.8	12.5±1.9	1.5±0.3

[a] Mean ± Standard Deviation

IV. Conclusion

Finger tapping test has been used in the evaluation of the muscle control and motor ability in the upper extremities especially for neurological disturbances such as in Parkinson's disease. Advanced setups allow the acquisition of finger-tapping signals that may reveal the test with more details and that may make it possible to assess the response duration time, response initiation time, and response rate from the timings for finger taps. However, sophisticated methods are needed to analyze the signals. Poincare analysis offers a method to analyze nonlinear and nonstationary signals in the time domain revealing three parameters to reflect the variability in the signals: *SD*1, *SD*2, and their ratio (ie. *SD*1/*SD*2). The method has been well-accepted to characterize the variability in heartbeat from the EGC signals [12].

In the current study, we aim to implement Poincare analysis using an embedded system to quantify the variability in finger-tapping signals in real-time. For the purpose, numerous finger-tapping signals are generated considering different levels of variation in response initiation time, and the signals are fed to a microcontroller unit. The microcontroller performs estimations for the response initiation time from the input signal, applies Poincare analysis on the estimates, and calculates the Poincare parameters. Results show that the microcontroller performs quite fast and outputs accurate values for the Poincare parameters making its use possible for real-time applications aiming quantification of variability in finger-tapping.

There are some limitations to the current study. All implementations and analyses are performed by simulations and the simulation environment may manage to model the practice with a certain degree that may lead to outstanding performances for the microcontroller unit. The finger tapping signal consists of a sequence of noise-free rectangular signal pulses for onsets of finger taps. Such a pulse may not be delivered by the sensing setup in practice and this may lead to some error in the response initiation time estimates resulting in poor accuracy in the Poincare parameters by the microcontroller. Poincare analysis is performed only for a set of twenty-five response initiation time estimates from the finger-tapping signal. Consideration of such a less number of estimates for the analysis may not figure out the long-term variability adequately and the Poincare parameter, SD2 may

not be generalized well. The microcontroller runs on 16MHz clock speed and reveals adequately fast Poincare analyses; however, a faster microcontroller would perform faster. On the other hand, there are some future improvements in the current work. The finger tapping variability has been quantified using response initiation time and we plan to expand our analysis for response duration time and response rate, the two other important measures for the quantification. We also plan to convert our work into practice when the covid-19 outbreak has lost its impact.

In conclusion, real-time calculation of Poincare parameters by embedded systems is feasible and offers simultaneous quantification of finger-tapping variability that opens up new opportunities for therapeutic devices. Additional work is needed to figure out the potential benefits in practice.

REFERENCES

[1] M S Shourijeh, R S Razavian, and J McPhe, "Estimation of maximum finger tapping frequency using musculoskeletal dynamic simulations," *J. Comput. Nonlinear Dynam.*, vol 12 pp 051009, Sep 2017

[2] D R Roalf, P Rupert, D Mechanic-Hamilton, L Brennan, J E Duda, D Weintraub, J Q Trojanowski, D Wolk, P J Moberg, "Quantitative assessment of finger tapping characteristics in mild cognitive impairment, Alzheimer's disease, and Parkinson's disease," *J Neurol* ; vol 265, pp 1365-1375, Jun 2018

[3] L Schmitt, "Finger-Tapping Test," Encyclopedia of Autism Spectrum Disorders, pp 1296, 2013

[4] R Okuno, M Yokoe, K Fukawa, S Sakoda, and K Akazawa, "Measurement system of finger-tapping contact force for quantitative diagnosis of Parkinson's disease," *Annu Int Conf IEEE Eng Med Biol Soc* , pp 1354-7, 2007

[5] M Djurić-Jovičić, N S Jovičić, A Roby-Brami, M B Popović, V S Kostić, and A R Djordjević, "Quantification of finger-tapping angle based on wearable sensors," *Sensors (Basel)*, vol 25, no 17, pp 203, Jan 2017

[6] K Shima, T Tsuji, A Kandori, M Yokoe, and S Sakoda, "Quantitative evaluation of human finger tapping movements through magnetic measurements," *IEEE/ASME Transactions on Mechatronics*, vol 24, no 1, pp 186-196, Feb 2019

[7] C Barut, E Kiziltan, E Gelir, and F Köktürk, "Advanced analysis of finger-tapping performance: A preliminary study," *Balkan Medical Journal*, vol 30, no 2, pp 71-167, 2013

[8] T Aoki, and Y Fukuoka, "Finger tapping ability in healthy elderly and young adults," *Medicine and Science in Sports and Exercise*, vol 42, no 3, pp 449-455, Mar 2010

[9] W Liu, L Forrester, J Whitall, "A note on time-frequency analysis of finger tapping," *J Mot Behav.*, vol 38, pp 18-28, Jan 2006

[10] Avaliable: https://wiki dfrobot com/DFRobot_Capacitive_Touch_ Sensor_SKU_DFR0030

[11] A Golińska, "Poincaré plots in analysis of selected biomedical signals," Studies *in Logic, Grammar and Rhetoric*, vol 35, pp 117-127, 2013

[12] M Brennan, M Palaniswami, and P Kamen, "Do existing measures of Poincaré plot geometry reflect nonlinear features of heart rate variability*?," IEEE T Bio-Med Eng.,* vol 48, no 11, pp 1342-1346, 2001

Spectral Analysis of Alcoholic Data

O. AKGUN[1] and T.C. AKINCI[2]

[1] Marmara University, Dep. of Computer Eng. Istanbul/Turkey, oakgun@marmara.edu.tr
[2]Istanbul Technical University, Dep of Electrical Eng., Istanbul/Turkey, akincitc@itu.edu.tr

***Abstract* - Today**, alcoholism is an important health problem in the world. Alcohol addiction is a brain disease and causes many changes in the structure of the brain. Electroencephalography has an important place as a non-invasive electrophysiological method in the investigation of the biological aspect of alcohol dependence. In this study, EEG data from control subjects (non-alcohol dependent) and alcohol-dependent subjects were used. The changes of these data with time were obtained first, then amplitude and phase spectra were obtained. EEG signs of control and alcohol addicts were compared spectrally. Structural differences between these two groups are determined based on frequency.

***Keywords* – EEG, alcoholic data, spectral analysis, fourier transformation.**

I. Introduction

SCIENTISTS Scientists have known for many years that alcohol directly or indirectly interacts with various neurotransmitters in the nervous system. What causes this interaction is due to ethyl alcohol crossing the brain-blood barrier. Because ethyl alcohol is a lipid-soluble substance, so it can reach the brain [1,2].

Alcohol addiction is a chronic brain disease like other substance addictions. It has been proven in many academic studies that long-term and heavy alcohol use causes changes in the brain structure. Structural changes are particularly evident in the anterior part of the brain (frontal region), indicating that the frontal region is prone to the toxic effects of alcohol. Regional volume losses in other cortical-subcortical areas, increase in cerebrospinal fluid, increase in ventricular volumes, generally shrinkage in the brain constitute the changes reported in neuroimaging studies [3, 4].

Electroencephalography (EEG) is valuable as a non-invasive electrophysiological method in the investigation of the biological aspect of alcoholism [5]. EEG is an examination method in which the spontaneous electrical activity of the brain is recorded through electrodes. This examination reflects the current functional state of the brain rather than its structural features. Therefore, despite advances in structural imaging methods (such as CT, MRI), it still maintains its importance. Especially in clinical pictures without a pathological finding reflected in structural examination methods, the importance of EEG increases even more [6, 7, 8].

The amplitudes of the brain waves on the skull surface are 0-200 microvolt frequencies and 1-50 Hz or higher. Brain waves are irregular. Different waves occur depending on the situation. EEG waves in normal people; It is classified as, α, β, θ and Δ. The frequency ranges of these waves are given in Figure 1.

Alpha (α) waves; They are rhythmic waves with amplitudes of 50μV and frequency between 8-13 Hz. In the EEG data of calm, rested and young people, it disappears in deep sleep. Visual stimulation stops α waves suddenly. Extraversion (less introverted) is linked to creativity and mental work. It is one of the most important frequencies of the brain in learning and using the taught information.

Beta (β) waves; The waves of the person who is awake and focused on something else disappear and are replaced by β waves. Its frequency is 13-32 Hz, it has high frequency and low amplitude. They are not synchronous waves. When attention is directed to the outside world, beta activity is in a normal waking state.

Theta (θ) waves; frequency is in the range of 4-8 Hz. In a state of frustration and emotional stress, In brain degenerative disorders, lethargy, daydreaming, sleep or 'indirect' may indicate the first stage of dreaming/thinking. Theta is believed to reflect activity in the limbic system and hippocampal region.

Delta (Δ) waves; It includes all frequencies less than 4 Hz (0.5-4Hz). It occurs in very deep sleep, its amplitude 2 to 4 times greater than others.

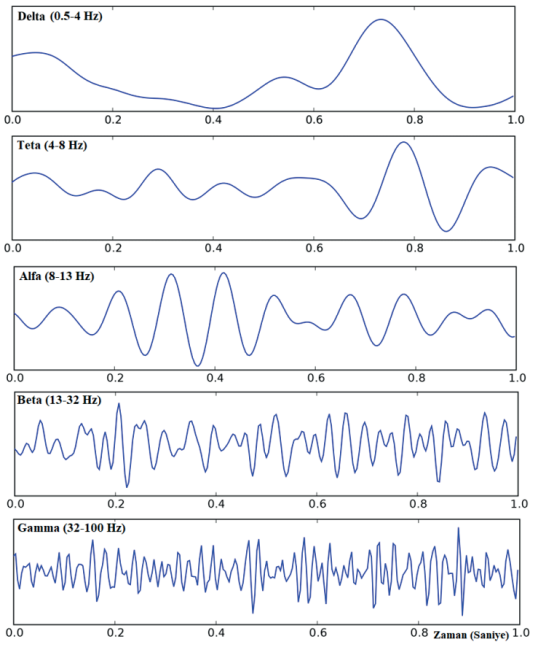


Figure 1. Brain waves and Frequency Bandwidth [9,10].

II. DATABASE AND DATA COLLECTION METHOD

In this study, EEG data were collected from the database. The collected data were grouped and analyzed as non-addicted (non-alcoholic) and alcohol addicted.

EEGs are recordings of the electrical activity of the brain with electrodes placed directly on the scalp or in the brain. 20 electrodes (EEG sensor) are placed on the scalp surface in standard positions adopted by the International Federation. This is called the 10/20 system (Figure 2) [9, 10,11].

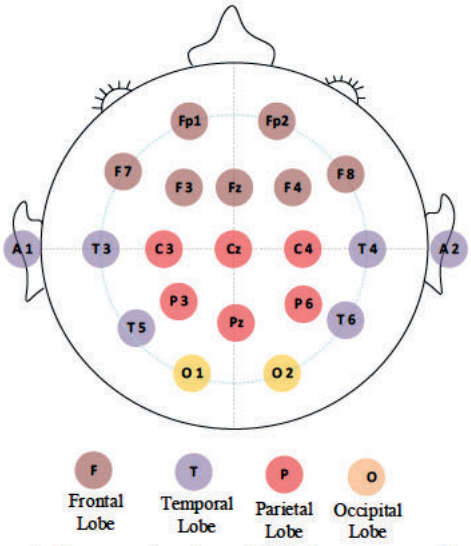


Figure 2. Placement locations of EEG electrons according to the 10/20 system [10].

As signal analysis methods developed, researchers working in this field were especially interested in biological signals such as EEG, EMG, and EKG. They tried to reach different information about signs by applying various analysis methods to these signals obtained from the body.

In this study, the characteristics of the models obtained by applying time-frequency and phase spectrum methods to EEG signals from normal and alcohol-dependent individuals were examined. The data used in the study were taken from The UCI KDD Archive [12]. 1-second recordings were obtained from 64 electrodes placed on the scalp of healthy and sick individuals at a sampling frequency of 256 Hz. Control and alcoholic subjects were exposed to a stimulus (a picture was shown to them) and recorded [12]. The time-amplitude graph of the signals obtained from the C4 electron in the study is given in Figure 3. The time-amplitude graph of the non-alcoholic (control) group is shown in Figure 3a, and the time-amplitude graph of an alcoholic individual is shown in Figure 3b.

In the figure, it is observed that the period of alcoholic EEG is significantly smaller than the control EEG and its amplitudes are mostly larger than those of the control.

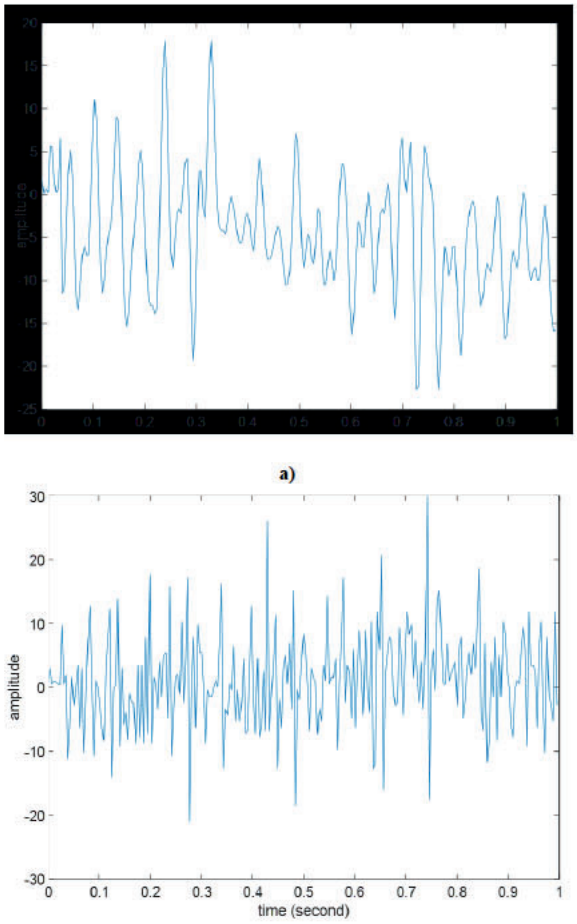


Figure 3. a) Control b) Change of alcoholic EEGs over time.

III. TIME FREQUENCY ANALYSIS AND MATHEMATICAL BACKGROUND

A. *Time-Frequency Analysis*

Many unknown events are hidden in an EEG signal recorded in time environment. Therefore, the signal alone is not sufficient in the time environment. It is necessary to pass to the frequency environment for analysis. Real and virtual component values in the frequency environment are obtained by taking the Fourier Transform of a signal recorded in time environment. Frequency-dependent amplitude and phase spectra are obtained with Fourier Transform. The amplitude spectrum shows the frequency content of an EEG signal. Phase spectrum determines the initial states of the waves of different frequencies that make up the EEG signal. Although their amplitude spectra are the same, there are seismic signals with different phase spectra. Therefore, looking at the amplitude spectrum of a seismic signal is not always sufficient. It also needs to examine the phase spectrum. Hence, finding a relationship between the "Argand diagrams" of the EEG signal spectrum, whose real component is the abscissa and the virtual component is the ordinate, is important in terms of defining the lithology of the environment where the EEG signal is recorded [13, 14, 15].

B. Mathematical Background

The Fourier Transform (FT) is not suitable for non-stationary signals. Using the windowing method, Gabor handled a small part of the sign in the time domain and expressed the sign in two dimensions as a function of time and frequency. In this transformation method, a certain part of the signal is passed through a window that can be considered to be stationary and FT operation is performed with a local frequency parameter. In Short Time Fourier Transform (STFT) the signal is divided into small frames and the signal is assumed to be stationary at these frame moments. These frames where stationarity is valid are obtained by multiplying the signal by a window function. This technique is based on the idea of localization of FT and the transformation process is performed by selecting an appropriate window in the place of interest. Spectral analysis of non-stationary signals requires skilled functions that must offer very good time and frequency resolution.

In the STFT method, if the current window centre value of the window function shifted on the time axis is τ, the windowed time series can be expressed with the following equation.

$$g(t) = f(t).g(t-\tau) \tag{1}$$

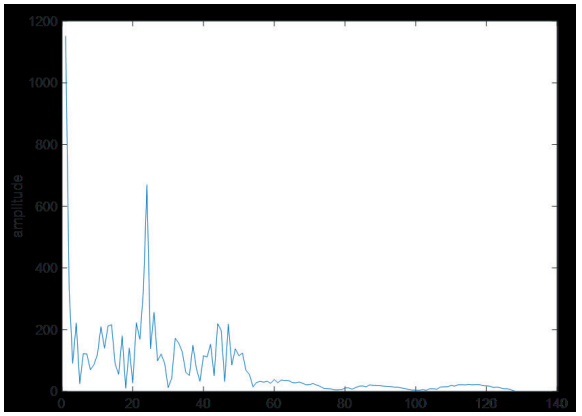
Then, equation (2) can be written if the Fourier Transform is taken for this respective time interval of the windowed time series given by equation (1).

$$KSFD(\tau, f) = \int_{-\infty}^{+\infty} g(t).exp(-i2\pi ft)dt \tag{2}$$

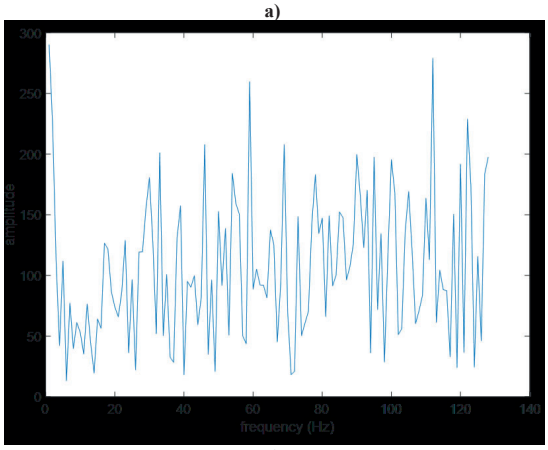
Then the window is shifted to a new time value and again the Fourier Transform of this windowed time series is taken. By continuing this process, a two-dimensional time-frequency image map can be created [16, 17, 18].

IV. Spectrum Analysis of Data

In this study, the EEG of two non-alcoholic and alcohol-addicted individuals was collected for one second with marked EEGs. Frequency-amplitude, frequency-phase spectrum, as well as three-dimensional spectrum analyzes, were performed from the collected data. In Figure 4a, the frequency amplitude analysis of the one-second EEG signal taken from a subject belonging to the non-alcoholic (control) group is given. In Figure 4b, the frequency-amplitude analysis of one-second data obtained from an alcoholic patient is given. Here, it is seen that the data of the non-alcoholic subject is limited to 50 Hz, while the data of the alcoholic subject oscillate at different amplitudes up to 130 Hz.



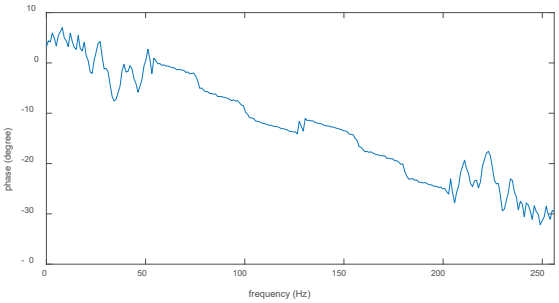
a)



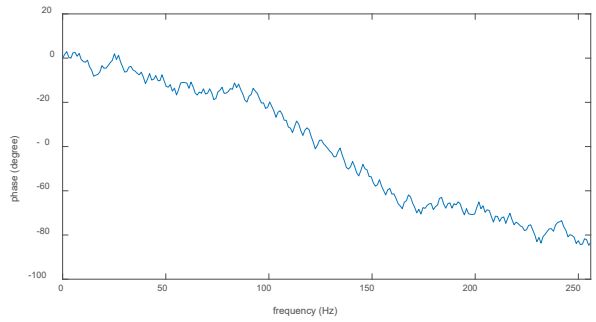
b)

Figure 4. a) Control b) Amplitude Spectra of Alcoholic EEGs.

While the amplitude spectrum of the control EEG shows a two-peak distribution with a bandwidth of approximately 50 Hz, the amplitude spectrum of Alcoholic EEG shows a multi-peak distribution distributed over the entire frequency plane (Figure 4).



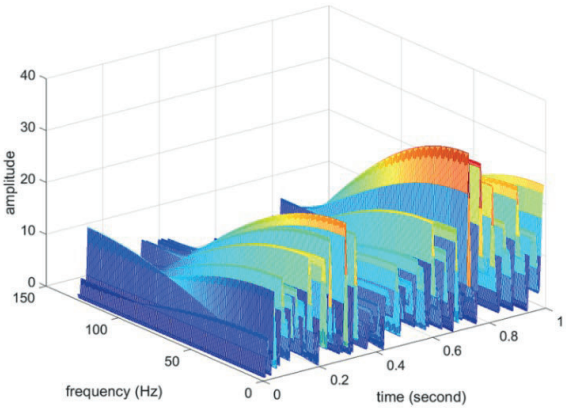
a)



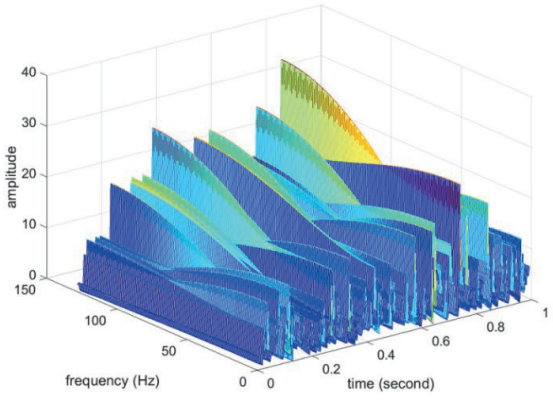
b)

Figure 5. a) Control b) Phase spectrum of alcoholic EEG.

In Figure 5, the Control EEG draws a phase curve that goes down to -30 degrees, while a phase spectrum descending to about -90 degrees is observed in the EEG data of the alcoholic subject.



a)



b)

Figure 6. a) Control b) Spectrogram of alcoholic EEG.

In Figure 6b, in the spectrogram of alcoholic EEG, especially the distribution scattered over the whole plane and showing high amplitude at high frequencies draws attention. Again, oscillations in smaller periods than in control are observed in alcoholic EEG.

V. Conclusions

The direct effect of alcohol on the brain is due to the fact that ethyl alcohol easily crosses the blood barrier of the brain. Long-term alcohol use causes structural changes in the frontal region of the brain. Regional volume losses in other cortical-subcortical areas, increase in cerebrospinal fluid, increase in ventricular volumes, and generally shrinkage in the brain are also observed. These structural disorders cause changes in the EEG, which measures the electrical activity of the brain. These changes have been observed in time, amplitude and phase spectra. In the change of EEG over time, the radiography of the alcoholic subject has a much lower period, but higher amplitude compared to that of the control subject. In the amplitude spectrum of the EEG, a two-peaked spectrum with a bandwidth of 50 Hz is observed in the control, while the spectrum in the data of the alcoholic subject has spread to the whole plane and the amplitudes are slightly lower. In the phase spectrum of the EEG, the phase of the control goes down to -30 degrees, while the phase of the alcoholic subject goes down to -90 degrees. With this analysis, this addiction can be easily understood by spectrum analysis of EEG data of two people who are not alcohol addicts and alcohol addicts. With this study, alpha, beta and delta waves of alcohol-addicted individuals can be examined separately and the distribution of these waves in these two groups can be determined.

References

[1] https://aklinizikesfedin com/alkol-bagimliligi-ve-norobiyoloji/19 şubat, 2020

[2] R Fama, A P Le Berre, S A Sassoon, N M Zahr, K M Pohl, A Pfefferbaum, E V Sullivan, “Relations between cognitive and motor deficits and regional brain volumes in individuals with alcoholism”, *Brain Structure and Function*, vol 224(6), pp 2087-2101; 2019

[3] K Sawyer, S Kayle, M O Berman, O J Barthelemy, G M Papadimitriou, G J Harris, N Makris, “Gender dimorphism of brain reward system volumes in alcoholism”, *Psychiatry Research: Neuroimaging* vol 263, pp 15-25, 2017

[4] N Koçuk, “Alkolün Beyin Yapısı Ve Bilişsel İşlevler Üzerine Etkileri”, Dokuz Eylül Üniversitesi Tıp Fakültesi Ruh Sağlığı Ve Hastalıkları Anabilim Dalı, Uzmanlık Tezi, 2010

[5] M F KARAASLAN, F O ORHAN, “EEG and ERP ın Alcoholism”, *Turkiye Klinikleri J Int Med Sci* , vol 1(47) pp 18-27, 2005

[6] B Baykan, E Altındağ, http://www itfnoroloji org/semi2/eeg htm, 2019

[7] M Durarıaj, R Jansı Ranı, “Mining and Extracting Emotional Information from Human-Based EEG: An Overview”, *Productivity*, vol 59 Issue 4, pp 343-351, 2019

[8] S Barra, M Fraschini, A Casanova, A Castiglione, G Fenu, “PhysioUnicaDB: a dataset of EEG and ECG simultaneously acquired”, *Pattern Recognition Letters*, vol 126, pp 119-22, 2019

[9] A Aykaç, EEG Elektroensefalogram, YDÜ Tıp Fakültesi Biyofizik AD, 2015

[10] A Özen, T Ç Akıncı, D Türkpençe, R Çağlar, S Şeker, “Beyin Dalgası Sensörü Kontrolü ile Akıllı Ev Otomasyonu”, EEMKON 2019, Elektrik Elektronik Mühendisliği Kongresi, 14-16 Kasım 2019, Istanbul-Türkiye

[11] F K Onay, C Kose, “Power spectral density analysis in alfa, beta and gamma frequency bands for classification of motor EEG signals”, 27th Signal Processing and Communications Applications Conference, SIU 2019, April 2019

[12] http://kdd ics uci edu/databases/eeg/eeg data html

[13] Y Duman, Z Kamacı, “Sismik Arazi Verileri Üzerinde Spektrum Çalışmaları Ve Litolojik Yaklaşımlar”, *DEÜ Mühendislik Fakültesi Fen ve Mühendislik Dergisi* vol 5 Sayı, 129-139 Ekim, 2003

[14] P Lindstedt, T Sudakowski, E Rokicki, “Machine Diagnosis Based on Amplitude-Phase Characteristics, Determined from the Experimental Amplitude Spectrum and the Calculated Phase Spectrum”, 2020

International Conference Mechatronic Systems and Materials (MSM) Mechatronics Systems and Materials (MSM), 2020 :1-4 Jul, 2020;
[15] B Tang, X Yang, H Wu, W Peng, "Forward-looking airborne radar clutter spectrum with amplitude and phase errors", Circuits and Systems (ICCCAS), 2013 International Conference on 1:276-279 Nov, 2013;
[16] https://acikders ankara edu tr
[17] T Bendory, D Edidin, Y C Eldar, "Blind Phaseless Short-Time Fourier Transform Recovery", *IEEE Transactions on Information Theory*, vol 66(5), pp 3232-3241, 2020
[18] M Sameer, A K Gupta, C Chakraborty, B Gupta, "Epileptical Seizure Detection: Performance analysis of gamma band in EEG signal Using Short-Time Fourier Transform", Wireless Personal Multimedia Communications (WPMC), 2019 22nd International Symposium on :1-6 Nov, 2019;

Wrist Print Region Segmentation Based on Deep Neural Networks

H. E. KOÇER[1] and K. K. ÇEVİK[2]

[1] Selcuk University, Konya/Turkey, ekocer@selcuk.edu.tr
[2]Akdeniz University, Antalya/Turkey, kcevik@akdeniz.edu.tr

***Abstract* - In recent years, biometric recognition based systems have become widespread. One of these is wrist-based recognition systems. In this study, wrist print based recognition system was developed by using near infrared (NIR) camera. Totally 220 NIR camera images taken from 10 for each both hands of 11 people. The obtained data set is allocated 70%(153 images) for training and 30%(67 images) for testing. The wrist regions are labeled on the training set images. The labeled data was trained with YOLOV2 architecture supported by ResNet50 one of the deep neural network models. The trained model was tested with the remaining 30% of the data set. In the test process, the wrist region was determined in the NIR images with the trained model. As a results of the study, it was seen that the wrist regions were correctly detected in all test images and the mean value of obtained similarity rates was %95.26. Therefore, it can be said that the deep learning architectures ResNet and YOLO are effective in the segmentation of the wrist region.**

***Keywords* - Wrist Print Recognition, Deep Neural Networks, Near-Infrared Camera, YOLO**

I. Introduction

BIOMETRIC recognition based systems are known as most secure systems. What makes biometric systems important in this way is that the security parameter they use is personal and there is no risk of theft and cannot be copied. Although there are many biometric recognition methods, fingerprint recognition, vein recognition and face recognition are mostly used today.

Vein recognition is a high-security biometric recognition approach based on human vascular structure. In this biometric identification approach, finger veins, hand veins, palm veins and wrist veins are used as biometric parameters. The identification systems which use wrist vein structure as biometric input are called wrist print recognition. In wrist print recognition system, the human wrist image is captured by using near-infrared (NIR) camera and illumination and then the wrist region is segmented for determining the wrist vein information. Then, the wrist print is used for identification.

In this study, the wrist vein region was detected and marked from the hand and wrist image taken using a near infrared camera. In this context, a software based on deep learning has been developed including ResNet50 and YOLO algorithms. The performance of the system was evaluated comparatively by testing the segmentation process on 220 wrist images taken from 11 people.

II. Deep Neural Networks (DNN)

Deep neural networks (DNN) can recognize objects without being affected by different properties such as different positions, directions and camera angles and environmental factors such as lighting. A deep learning algorithms are trained on tagged images. An architecture of DNN is shown in Figure 1.

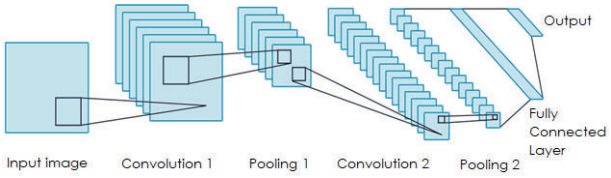


Figure 1: A Deep neural networks architecture.

The input usually consists of images or signals. In the convolution layers, the filtering process is applied to the 3-dimensional matrices in the previous layer. The number of filters used constitutes the depth (the size of the 3rd dimension) of the convolution layer. In the pooling layer, size reduction is applied. In the section called fully connected layer, classical artificial neural network operations are performed. The output can be defined as a vector with the length of the defined class.

The number of layers in the architecture and the filter size and number can be changed by the user to suit the application. In addition, the performance may vary depending on the number of images and iterations to be used for training. If the number of training iterations is too high, both the training time increases and the model moves towards memorization. As the rate of misclassification in data that the model has not seen before will increase as a result of memorization, accuracy decreases. In the study, it was tried to determine the wrist region from the images taken by using the ResNet50 and YOLO architectures.

A. ResNet50

The Residual Network (also known as ResNet) uses redundant blocks with multiple layers to reduce training error. Typical ResNet models are implemented with double- or triple-layer skips that contain nonlinearities (ReLU) and batch normalization in between. During training, the weights adapt to mute the upstream layer, and amplify the previously-skipped layer. In the context of residual neural networks, a non-residual network may be described as a plain network [1, 2].

The standard ResNet architecture consists of 152 layers and has a higher classification success than other DNN architectures. The ResNet50 architecture used in this work, on the other hand, consists of 34 layers. Both layer blocks in the network are replaced by a three-layered residual block. The residual block characterization can be given as:[3, 4]

$$a^{[l+2]} = g(a^{[l]} + z^{[l+2]} \quad (1)$$

Here, a is the activations (outputs) of neurons in layer l, g is the activation function for layer l. In this equation, l is used for clarity.

B. You Only Look Once (YOLO)

YOLO is an algorithm for object detection using convolutional neural networks (CNN). It stands for "You Only Look Once", which means "Just Look Once". The reason for choosing this name is that the algorithm is fast enough to detect objects in one go. The most important feature that distinguishes YOLO from other algorithms is its ability to detect objects in real time. YOLO uses Darknet library and has a very fast running architecture. The reason it is so fast is that it can guess objects and coordinates by passing any picture through the neural network in one go. In other words, YOLO takes an approach that solves a single regression problem while performing the estimation process [5, 6].

When we look at YOLO's algorithm, we first see that it divides the input picture into grids of certain sizes (3x3, 5x5, 17x17 etc.). It is then checked whether the searched object is within the grid. The object detected in the similarity check is scored and marked. Then non maximum suppression is applied, which evaluates the score of each marked object, and those below a certain confidence score are separated [7, 8].

C. Related Works

Although many studies have not been done on wrist vein recognition, we come across studies using different approaches in the literature. Wang et al. (2007), images of both wrists were taken with an 800-850 nm NIR camera for biometric purposes [9]. They were applied Median & 2D Gaussian filter for noise reduction [10]. For feature matching, Hausdorff Distance method was used in this study [11]. Uriarte-Antonio et al. (2011) aimed to recognize the right hand wrist veins using the minutiae extraction approach [12]. Vector minutiae comparison is used in feature matching process. In the same year, Hartung et al. performed spectral minutiae extraction in the identification of wrist vessels [13]. In preprocessing stage, binarization and Skeletonization were implemented after noise reduction and edge enhancing. Convolution approach and Location-Based Spectral Minutiae Representation were used for extracting the feature. In this study, Hausdorff distance, Modified Hausdorff, Similarity-based Mix-matching, SML correlation and SML fast rotate methods were used for feature matching [14]. Das et al. In their study in 2014, they performed the feature matching process with Support Vector Machines [15, 16]. Dense Local Binary Pattern approach was applied for extracting the feature. In 2018, Nikisins et al. studied wrist vein recognition with cross-correlation-based comparison [17]. Adaptive histogram equalization [18] and Discrete Meyer Wavelet techniques [19] were used in preprocessing phase. Then Hessian matrix was implemented and the wrist vein features were extracted in this study. In 2020, Garcia-Martin et al. were developed portable contactless system for wrist vascular biometric identification. a new contactless database has been collected with the software algorithm TGS-CVBR®. In this work, Preprocessing and Identification Software for Contactless Vascular Biometric Recognition (PIS-CVBR®), is the recognition algorithm [20].

III. Suggested Method

The flow chart of the system designed within the scope of the study is shown in Figure 2.

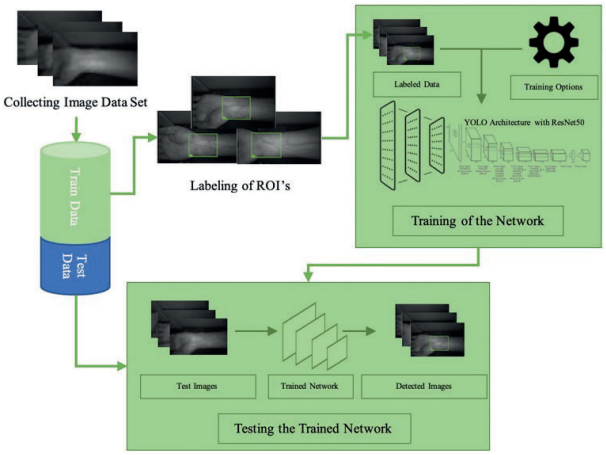


Figure 2: Flow chart of the designed system.

The study is primarily provided within the scope of the data collection. Two wrist images of 11 people were taken with a near infrared (NIR) camera. In the study, a NIR camera with a resolution of 752x480 pixels, a monochrome (gray level) and an infrared light source operating at 880nm wavelength was used to capture the images. By taking 10 images for both wrists of each person, a total of 220 images were provided in the data set. The images obtained are gray level with a resolution of 752x480 pixels in RGB format. Images consist of 8 bits as color depth. The sample images used in the designed system are shown in Figure 3.

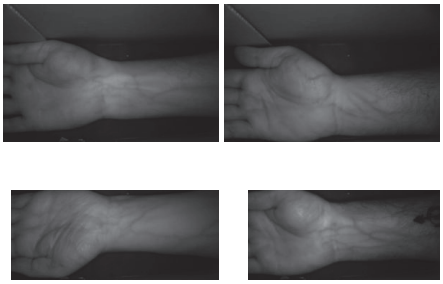


Figure 3: Image set examples.

In the second stage, the data set is divided into training and test data. As used frequently in the literature, the obtained images were determined as homogeneous 70% (153 images) for training and the remaining 30% for testing.

In the next stage, Region of Interest (ROI) areas are labeled on the images (70%) allocated for training. In order to define the wrist print region, ROI is labeled as 224x224 dimensioned boxes where the veins on the wrist are clearly visible. Image Labeler plugin of Matlab software was used for this labeling process. An example of image labelled in the training set is shown in Figure 4. The remaining test data were transferred to the test stage without labeling.

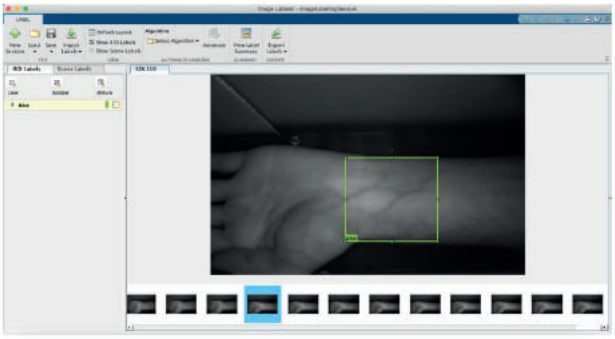


Figure 4: Image labeling process.

After the data labeled and prepared, the labeled data was trained with Deep Neural Network - (DNN) with the software realized within the scope of the study. In the training phase of the network, the images labelled were given for 1 class (wrist) of output as an introduction to the neural network built from YOLO architecture powered by ResNet50 and training has been done. Layer structure of the designed network is shown in Figure 5.

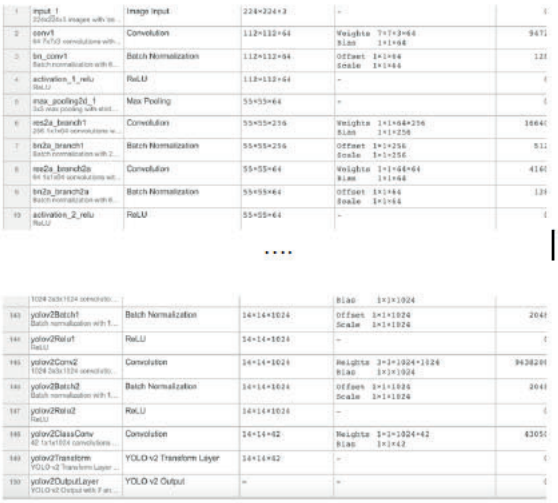


Figure 5. YOLO architectures with ResNet50

As can be seen in Figure 5, ResNet50 was created in a 150-layered structure because of integration with YOLO. Since the input data of the ResNet network is **224x224x3**, the images in the data set have been adapted to this resolution. The single class "wrist value" is given as output data. The designed deep neural network was subjected to the training process with the parameters given in Table 1.

Table 1: Training parameters for proposed method.

Option	Value
Epoch (Iteration)	100 (1300)
Batch Size	16
Rate	0,001
Optimizer	SGDM
Input Size	224x224x3
Output Size	1 (Wrist)
Verbose	True
Verbose	1
Shuffle	Never

The computer on which the designed model is trained has the following hardware configuration: Intel i7-2600 3.40 GHz processor running under 64 Bit Windows 10 operating system, 22 GB RAM, GTS450 GDDR5 1GB 128Bit Nvidia GeForce DX11 Graphics Card hardware. In addition, Matlab software was used in coding, training and testing of the model.

The change of Root Mean Square Error (RMSE the square root of the difference between the value obtained and the value obtained during the training of the network) and LOSS (expressing how different the estimate made by the model from the actual value) values are given in Figure 6. In addition, the training time of the network is 03:46:10, the resulting minimum RMSE value of the network is 0.07 and LOSS value is 5.2e-03.

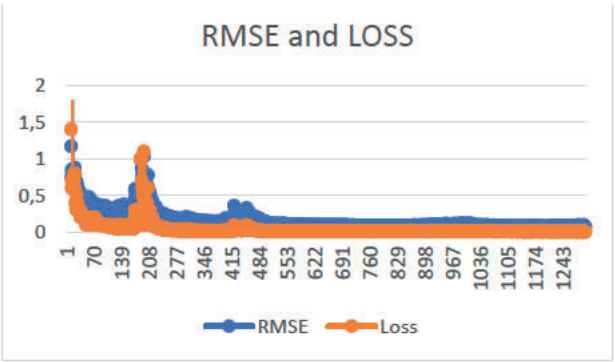


Figure 6: Change of RMSE and LOSS values during training

IV. Results and Discussion

For the evaluation of the study, the test process was carried out by using images (30%) that were not shown to the network model whose training was completed. Test images were given as an input to the network and the wrist areas were determined. In determining the region, the minimum value of similarity ratio was taken as 0.70 and the highest similarity value was determined as the result. Sample test images including the detected and marked boxes are given in Figure 7.

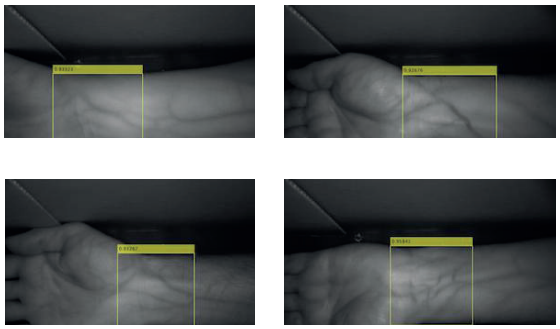


Figure 7: Detection of the wrist area

The percentage values for determining the wrist area, which are formed by input (total of 67 images in the test process), are shown in Figure 8. The process of determining the wrist area on an image takes approximately 0.42 seconds. It took approximately 30 seconds to determine the ROI areas of all test images and the ROI's were saved as a file.

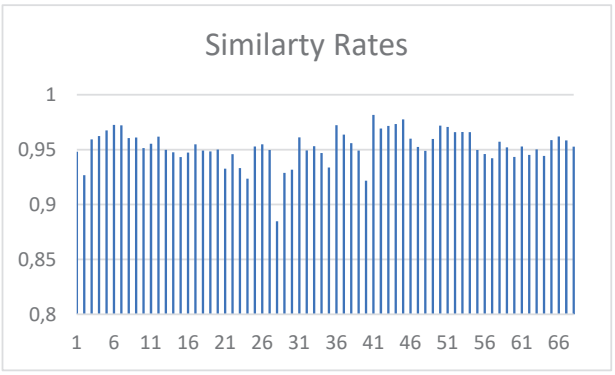


Figure 8: Similarity rates for test images

As a result of the study, it was seen that the wrist regions were detected correctly in all test images and the average value of the obtained similarity rates was 95.26%. Therefore, it can be said that the deep learning architectures ResNet50 and YOLO are effective in the segmentation of the wrist region.

REFERENCES

[1] Wikipedia "Residual neural network " Wikipedia https://en wikipedia org/wiki/Residual_neural_network (accessed 15 11 2020, 2020)

[2] K He, X Zhang, S Ren, and J Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp 770-778

[3] A Amidi, S Amidi, D Vlachakis, V Megalooikonomou, N Paragios, and E I Zacharaki, "EnzyNet: enzyme classification using 3D convolutional neural networks on spatial representation," *PeerJ,* vol 6, p e4750, 2018

[4] S Amidi, A Amidi, D Vlachakis, N Paragios, and E I Zacharaki, "Automatic single-and multi-label enzymatic function prediction by machine learning," *PeerJ,* vol 5, p e3095, 2017

[5] B Kuyumcu, B Buluz, and Y Kömeçoğlu, "Author Identification in Turkish Documents with Ridge Regression Analysis," in *2019 27th Signal Processing and Communications Applications Conference (SIU)*, 2019: IEEE, pp 1-4

[6] J Redmon and A Farhadi, "YOLO9000: better, faster, stronger," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp 7263-7271

[7] C C Nguyen *et al.*, "Towards real-time smile detection based on faster region convolutional neural network," in *2018 1st International Conference on Multimedia Analysis and Pattern Recognition (MAPR)*, 2018: IEEE, pp 1-6

[8] J Redmon, S Divvala, R Girshick, and A Farhadi, "You only look once: Unified, real-time object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp 779-788

[9] L Wang, G Leedham, and S -Y Cho, "Infrared imaging of hand vein patterns for biometric purposes," *IET computer vision,* vol 1, no 3, pp 113-122, 2007

[10] L Hong, Y Wan, and A Jain, "Fingerprint image enhancement: algorithm and performance evaluation," *IEEE transactions on pattern analysis and machine intelligence,* vol 20, no 8, pp 777-789, 1998

[11] M -P Dubuisson and A K Jain, "A modified Hausdorff distance for object matching," in *Proceedings of 12th international conference on pattern recognition*, 1994, vol 1: IEEE, pp 566-568

[12] J Uriarte-Antonio, D Hartung, J E S Pascual, and R Sanchez-Reillo, "Vascular biometrics based on a minutiae extraction approach," in *2011 Carnahan Conference on Security Technology*, 2011: IEEE, pp 1-7

[13] D Hartung, M A Olsen, H Xu, and C Busch, "Spectral minutiae for vein pattern recognition," in *2011 International Joint Conference on Biometrics (IJCB)*, 2011: IEEE, pp 1-7

[14] H Chen, G Lu, and R Wang, "A new palm vein matching method based on ICP algorithm," in *Proceedings of the 2nd International Conference on Interaction Sciences: Information Technology, Culture and Human*, 2009, pp 1207-1211

[15] A Das, U Pal, M A F Ballester, and M Blumenstein, "A new wrist vein biometric system," in *2014 IEEE Symposium on Computational Intelligence in Biometrics and Identity Management (CIBIM)*, 2014: IEEE, pp 68-75

[16] W H Press, S A Teukolsky, B P Flannery, and W T Vetterling, *Numerical recipes in Fortran 77: volume 1, volume 1 of Fortran numerical recipes: the art of scientific computing* Cambridge university press, 1992

[17] O Nikisins, T Eglitis, A Anjos, and S Marcel, "Fast cross-correlation based wrist vein recognition algorithm with rotation and translation compensation," in *2018 International Workshop on Biometrics and Forensics (IWBF)*, 2018: IEEE, pp 1-7

[18] S M Pizer *et al.*, "Adaptive histogram equalization and its variations," *Computer vision, graphics, and image processing,* vol 39, no 3, pp 355-368, 1987

[19] I Daubechies, *Ten lectures on wavelets* SIAM, 1992

[20] R Garcia-Martin and R Sanchez-Reillo, "Wrist Vascular Biometric Recognition Using a Portable Contactless System," *Sensors,* vol 20, no 5, p 1469, 2020

A Novel Approach and Application of Time Series to Image Transformation Methods on Classification of Underwater Objects

A. CIVRIZOGLU BUZ[1], M. UMUT DEMIREZEN[2] and U. YAVANOGLU[3]

[1] HAVELSAN, Ankara/Turkey, acivrizoglu@havelsan.com.tr
[2] ROKETSAN, Ankara/Turkey, umut@demirezen.tech
[2] Gazi University, Ankara/Turkey, uraz@gazi.edu.tr

Abstract **- Sonar is used to determine the size, distance, direction, and other object features using sound waves. It is widely used in submarine mining, oil exploration, submarine mapping, tracking fish shoals, and mine detection. The advanced engineering technology has led to the development of underwater platforms such as developing manned and unmanned platforms for anti-mine action operations in the military field and increased efficiency of submarine oil and mineral exploration. In addition to mines, other objects that are very similar in shape and structure to mines can be observed in the submarine. Recently detecting, identifying, and classifying sonar signals are topics that are studied widely. Feature extraction, feature selection, selection of the most appropriate algorithms, and hyper-parameter optimization studies of these algorithms, which should be used to identify and classify sonar signals, are seen as scientific problems studied for many years. In this study, instead of commonly used machine learning algorithms and feature extraction processes, it is suggested to use mathematical conversions with an innovative approach. The performances of deep learning methods are compared in this problem. With an innovative approach applied on a data set containing 208 signal information in time-series format; Time series data has been transformed from one-dimensional data to two-dimensional format by using conversion methods with specific features to image format. The conversion methods' performance was measured by the classification results of mines and rocks using deep learning algorithms on the pictures obtained. Moreover, the performance results obtained with deep learning algorithms and comparing the results obtained with classical algorithms are analyzed in detail. Compared with other studies in the literature, it has been recognized that the proposed time series data-to-image approach eliminates the need for feature extraction and obtains sufficiently good results.**

Keywords **– MTF, GAF, RP, CNN, Sonar, Deep Learning. Support Vector Machine, Machine Learning.**

I. INTRODUCTION

In recent years, exploration of the underwater world, detection and classification of underwater objects and targets occupy an essential place in many applications for civil and military purposes. Sonar (Sound Navigation and Ranging), which is used to explore the underwater world, is used to gain information about the size, distance, and structure of objects using sound waves [1]. Although sonar systems were initially developed for naval warfare, today they are used in many essential areas with the purpose of oil and mineral exploration, seafloor mapping, detection of fish shoals, measurement of waves and currents, detection of sunken ships, drowned and submerged people, wreckage such as aircraft and underwater search and rescue purposes in many essential areas [2]. Today, sonar systems occupy a vital place in the defense industry for underwater target and mine detection [3]. Not only in times of war, but especially in times of peace, often used to protect coastal security, ship landings, and port controls, and such matters as conducting sonar systems healthily, have started to offer more advanced opportunities in the fields with the development of technology [4]. With the development of the technology, both the quality and the features of the signals, and images received from sonar devices have increased. In recent years, although the lack of data sets and resources, especially for mines and similar objects, several studies conducted with both images and signals has made research in this field difficult, still there are a limited number of studies based on machine learning-based and statistical analysis methods published in the literature [4], [5], [6], and [7].

Detection of mine objects with artificial intelligence algorithms becomes difficult due to the high diversity of underwater mines, the similarity of other objects such as rocks, debris, and cables to mines in shape and structure, signal reflections, and disruptive noise effects on the signal. Classical machine learning algorithms are based on feature extraction. Obtaining these attributes explicitly developed for the problem can be done with field expertise. Besides, choosing the most important features among the produced features plays a vital role in training machine learning algorithms and developing and improving the results. As a result of the increase in the number of features named "curse of dimension" [8], generating an excessive number of features causes an exponential increase in the number of training data required for the algorithm. Using less training data and too many attributes cause the algorithm to overfit. When such situations arise, machine learning algorithms get unsuccessful results in classification or regression problems because they lose their ability to generalize. Although deep learning methods, in which feature extraction and selection processes can be performed automatically instead of classical algorithms, provide more advantages in this regard, they require more data to achieve

successful results. This study proposes a new method for the binary classification problem using a dataset containing sonar signal data obtained using cylindrical mines and cylindrical rocks similar in shape and size to mines [9]. Using the proposed method, we focused on the classification problem of rocks and mines, a dual classification problem, and compared classical machine learning algorithms and deep learning algorithms. Compared to other literature studies, converting one-dimensional time-series data to image format with an RGB channel was used. Changing the data representation ensured that convolutional neural networks-based deep learning architectures were used, and it was determined that successful results were obtained with limited data. In this paper; in the second section, a brief literature review summarizes the algorithms used and examines other studies using the dataset used in this paper. In the third section, the mathematical model and the technical explanation of the innovative approach used in this study are made. In the last section, all the results for experiments performed in order to compare algorithm performances are shown, and the findings are discussed.

II. Literature Review / Related Work

A study conducted in the field of classification of underwater objects [9] determines whether the target is a cylindrical mine or a cylindrical rock according to the returning signal from two different underwater objects. In this study, to compare the effects of the number of hidden layers and the viewing angle on the result, experiments have been made in different hidden layers and depending on and not depending on the viewing angle. K-Nearest Neighbor Classification algorithm (K-NN) was used in their studies. The performance of the people who examined these signals with the latest results was compared, and it was found that human performance remained low compared to the studies. The resulting high accuracy value of 90.4% depends on the viewing angle and obtained in trials with 12 hidden layers. In another study performed with the same data set [10], the Generalized Regression Neural Network (GRNN) method was used to classify the sonar data set. When using the GRSA method, when they added the feature extraction method known as Principal Component Analysis (PCA) to the study, it was observed that the classification performance improved.

In this study, the accuracy of the classification was evaluated by ROC analysis. The result was 91.34% when only the GRSA method was used, while the result was 93.26% when TBA was used in addition to GRSA. In another study conducted on this problem [11], nine sample-based genetic algorithms (GA) classifiers were used in various variations to find a better performance in generalization ability and training efficiency. The performance obtained as a result of these variations was compared. The nine methods mentioned have been tested with 19 datasets. The sonar dataset's highest success rate was 53.43% in the model named TS2. In another study using the C4.5 decision tree [12], the problem of classifying sonar targets by improving the performance of the C4.5 decision tree by adding the Knowledge Acquisition (IG) feature evaluation process was studied. Another point worth noting in this study is that feature selection improves computational efficiency while increasing the classification accuracy. It has been stated that the accuracy of the C4.5 algorithm, which is working with BK, has increased by more than 10% and reached 81%. In another study conducted on the analysis of sonar signals [13], sonar signals will turn into a data set of increasing size with an uninterrupted data flow. The performance problems of this size have been investigated. It was emphasized that data processing time should be kept as short as possible to detect data with a continuous flow such as sonar signals in real-time. In this study, a new strategy called iDSM-CA, which is designed to clean noisy data gradually using fast collision analysis, is presented as a solution to this problem. With this strategy, sonar classification tests were performed with six classification algorithms in total. In another study conducted on a similar problem [14], it is aimed to solve the complexity of the decision boundary that algorithms have to find in the classification field.

In order to solve this problem, the Incremental Maximum Gaussian Mixture Partition (IMGMP) algorithm is proposed for classification. This algorithm uses a divide and conquers strategy to calculate a suitable decision boundary. K-Means algorithm was used to divide the data into clusters. It was emphasized that the performance of IMGMP is highly dependent on the cluster size in establishing the decision boundary, and a result with an accuracy of 82.79% was obtained. In another study [15], a new algorithm is proposed to solve this problem of the nearest neighbor classification K-Nearest Neighbor Classification Algorithm (KNN), which occurs because the calculation cost increases when the size of the data set increases. While the KNN algorithm calculates the distance between the training data and the test data, the MIME-KNN algorithm has been proposed, which will maximize the distance between classes and the minimum in-class distance due to the complexity of this calculation as the data set grows. As a result of the studies, while the KNN algorithm provides an accuracy rate of 74.50%, the accuracy rate of the MIME-KNN algorithm has been obtained as 82.20%.

The results of all these studies are given in Table I in the Results section.

III. Material and Methods

The proposed method is based on converting time series data to three-channel (RGB) or single-channel (gray) lossless image format. In this transformation, the spatial and temporal changes of data in each frequency band are handled differently, and transformed into a three-dimensional picture format by one-dimensional data transformation. In this newly-converted picture format, it is possible to create additional features that cannot be used or obtained in one dimension, and with this information representation results in increasing classification performance. When the time series data is converted into image representation, the classification problem can be solved using classical image processing methods on these pictures. Two hundred and eight different signals obtained from 60 different bands from the sonar device were broadcast in time-series format [9]. Data in this format is transformed into an image format (RGB) using transformations with specific properties. In

the study, the visual transformation of the time series data was carried out with three different methods called Gramian Angular Field, Markov Transition Field, and Recurrence Plot.

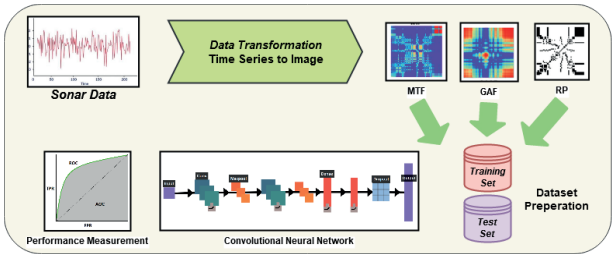


Figure 1: Flow chart for the proposed method.

A. Time Series-Image Transformation Approach

In this study, three different mathematical transformation methods are used to obtain images from the time series. These methods were used as Gramian Angular Field (GAF), Markov Transition Field (MTF) [16], [17], and Recurrence Plot (RP) [18], respectively. The RGB color map represents the values transformed by GAF and MTF methods. In the color map, colors close to blue and blue represent smaller values, while colors close to red and red represent the values with larger amplitudes. RP, on the other hand, is represented by a black and white image with a single channel. The time-series data properties are defined as 2-dimensional images with colors, dots, and lines at the corresponding positions in the image in which they are transformed and differ from each other.

1) Gramian Angular Field Transformation

In the angular field, the time series is represented in the polar coordinate plane instead of the Cartesian coordinate plane. In the GAF, each element is the cosine of the sum of the angles. GAF has two different methods. These methods are Gramian Angular Summation Field / GASF Gramian Angular Difference Field / GADF. In the Gramian angular area method, each point in the given time series is rescaled. So each value in Eq. (1) with [-1, 1] or in Eq. (2) with [0, 1].

$$\tilde{x}_{-1}^{i} = \frac{(x_i - max(X) + x_i - min(X))}{max(X) - min(X)} \tag{1}$$

$$\tilde{x}_{0}^{i} = \frac{x_i - min(X)}{max(X) - min(X)} \tag{2}$$

The obtained new time series values are expressed in the polar coordinate plane by encoding the time stamp as angular cosine, and the time stamp as a radius. In Eq. (3). Here $\emptyset$ angle is a fixed parameter for regulating the polar coordinate, t_i timestamp, r radius, and N polar coordinate plane.

$$\begin{cases} \emptyset = \arccos(\tilde{x}_i), -1 \le \tilde{x}_i \le 1, \tilde{x}_i \in \tilde{X} \\ r = \frac{t_i}{N}, t_i \in \mathbb{N} \end{cases} \tag{3}$$

$$GASF = \begin{bmatrix} \cos(\emptyset_1 + \emptyset_1) & \cdots & \cos(\emptyset_1 + \emptyset_n) \\ \vdots & \ddots & \vdots \\ \cos(\emptyset_n + \emptyset_1) & \cdots & \cos(\emptyset_n + \emptyset_n) \end{bmatrix} \tag{4}$$

$$GASF = \tilde{X}'.\tilde{X} - \sqrt{I - \tilde{X}^2}'.\sqrt{I - \tilde{X}^2} \tag{5}$$

After the time series are rescaled, they are transformed into a polar coordinate system with the help of Eq. (3). Using this data, converted into a polar coordinate system, according to the desired output, trigonometric addition or difference is transformed into matrix structure via Gramian Angular Summation Field (GASF) Eq. (4), Eq. (5) or Gramian Angular Difference Field (GADF) Eq. (6), Eq. (7). The conversion is completed by using the cosine functions for GASF and the sine functions for GADF.

$$GADF = \begin{bmatrix} \sin(\emptyset_1 - \emptyset_1) & \cdots & \sin(\emptyset_1 - \emptyset_n) \\ \vdots & \ddots & \vdots \\ \sin(\emptyset_n - \emptyset_1) & \cdots & \sin(\emptyset_n - \emptyset_n) \end{bmatrix} \tag{6}$$

$$GADF = \sqrt{I - \tilde{X}^2}' \cdot \tilde{X} - \tilde{X}' \cdot \sqrt{I - \tilde{X}^2} \tag{7}$$

At the end of this step, the data in the time series format is encoded in a picture format containing 3-channel RGB color coding, and the transformation process is completed. An application of the described method is shown in Figure 2.

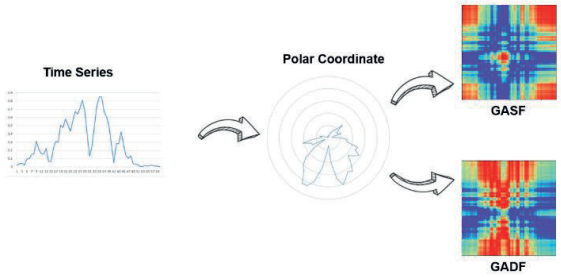


Figure 2: Gramian angular field transformation.

2) Markov Transition Field Transformation

Markov transition matrix with transition probabilities for the time series decomposed in MTF is created. This matrix is then transformed into an image. In the Markov Transition Field, a matrix of size $Q \times Q$ is created from a given time series. First of all, the partitioning process is completed by defining the time series' quantitative divisions and assigning each to its corresponding quantile divisions. In the equation given in Eq. (8), it shows the Markov Transition Matrix obtained by dividing the data by the Q quantile interval:

$$M = \begin{bmatrix} w_{ij|x_1 \in q_i, x_1 \in q_j} & \cdots & w_{ij|x_1 \in q_i, x_n \in q_j} \\ w_{ij|x_2 \in q_i, x_1 \in q_j} & \cdots & w_{ij|x_2 \in q_i, x_n \in q_j} \\ \vdots & \ddots & \vdots \\ w_{ij|x_n \in q_i, x_1 \in q_j} & \cdots & w_{ij|x_n \in q_i, x_n \in q_j} \end{bmatrix} \tag{8}$$

The transformation process is completed by encoding the probability-based matrix obtained with this transformation in a picture format that includes 3-channel RGB color coding. An application of the method described is shown in Figure 3.

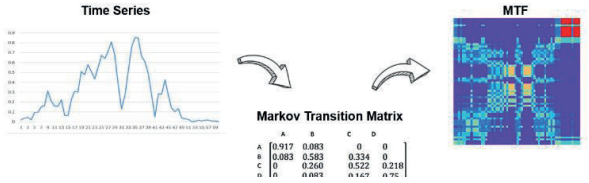


Figure 3: Markov transition field transformation.

3) Recurrence Plot Transformation

Recurrence Plot; it is a method used for repetitive processes such as time series, periodic data, and cycles and visualizes these processes in two dimensions. RP is also an advanced technique that is frequently used in the applications of nonlinear data analysis. The primary purpose of recurrence plots is to examine the higher dimensional phase space trajectories visually. RP's can be applied to short and unstable data. In the Recurrence Plot method, the image is created by calculating the distance of each point in the time series to the same time series. In this method, which works to visualize a time-dependent behavior, the primary step of the visualization is calculating the $N \times N$ dimensional matrix obtained from the time series [18], a tool that can visualize the repetition of x_i in i^{th} states in phase space. The mathematical representation of the repetition graph is shown in Eq. (9).

$$\mathbf{R}_{i,j} = \Theta\left(\varepsilon - \|\vec{x}_i - \vec{x}_j\|\right), i,j = 1 \dots N \tag{9}$$

In Eq. (13) $\mathcal{E}$ is a threshold, N is the number of data points to be calculated, Θ is a stepwise function. This tool allows to investigate the m-dimensional phase space trajectory through a two-dimensional representation of its repetitions. Such a repetition of a state at a different time is marked by ones and zeros (black and white points on the graph) within the two-dimensional square matrix. This representation is called the Recurrence Plot (RP). The transformation process is completed by encoding the bivalent matrix obtained in a picture format containing one channel color coding. An application of the method described is shown in Figure 4.

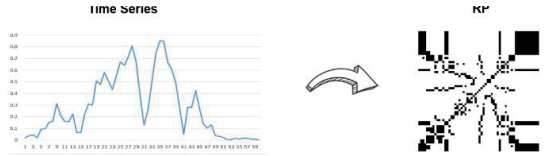


Figure 4: Recurrence plot transformation

IV. Dataset

The data set [9] used in this study was created with the signals reflected from cylindrical mines and cylindrical rocks resembling these mines from different angles. The cylindrical mines and cylindrical rocks from which the sonar data sets were obtained are approximately 5 ft. long. There are a totals of 208 data in the data set. These data were obtained by taking an average of 5 turns to be chosen randomly from each viewpoint out of 1200 signals sent as a return to signals sent from different angles, provided that they are up to 180 degrees for rocks and 90 degrees for mines. Of the 208 data obtained, 111 belong to cylindrical mines and 97 to cylindrical rocks. Each data is a set of 60 numbers, ranging from 0.0 to 1.0. Each number represents the energy collected over a fixed period in a given frequency band. Although each number does not precisely reflect the angle value, they are ordered in ascending angles. In the classification of each data, the data rock is labeled as R, and the mine as M. In Figure 5 and Figure 6, the graph of the sonar signal returns in the time series axis belonging to all mines and all rocks in the data set. As can be understood from these graphics, there are differences in returning to the sonar signals sent to objects.

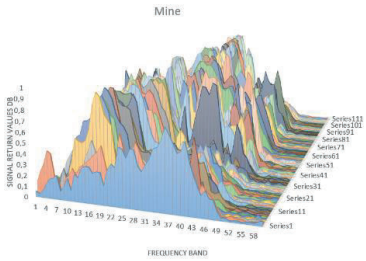


Figure 5: Signal return values corresponding to the frequency band measured for all mine targets.

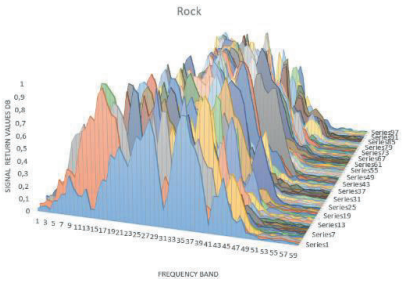


Figure 6: Signal return values corresponding to the frequency band measured for all rock targets.

V. Brief Description Of Used Machine Learning Methods

In this study, classical machine learning algorithms, which are frequently encountered in the literature, were selected for this binary classification problem. Selected classic machine learning algorithms are Support Vector Machine, Decision Tree Algorithm, K-Nearest Neighbor Algorithm, Linear Discriminant Analysis, Gauss NB, Logistic Regression, and Random Forest. Their performance was optimized by hyper-parameter selection specific to the algorithm. In the following subsections, these methods are briefly described and compared:

A. Convolutional Neural Networks (CNN)

Convolutional neural network is a deep neural network consisting of neurons with learnable weights that can process data strings such as image classification, widely used in computer vision. As the proposed deep learning architecture, a

custom convolutional neural network given in Figure 7 has been designed, and used due to the small number of data. It has been determined that a less complicated network is appropriate because other architectures commonly used in the literature generally overfit and memorize the data.

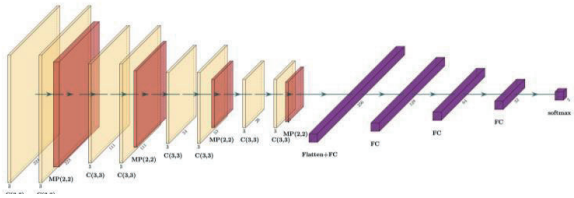


Figure 7: The convolutional neural network architecture used.

B. Support Vector Machines algorithm (SVM)

SVM is a supervised learning algorithm, is a vector space-based algorithm that determines hyperplanes to separate data from each other by finding a decision boundary between data. Among the advantages of the SVM algorithm are that it is effective in high-dimensional data, makes memory usage more efficient thanks to the training points it uses, and makes versatile decisions with different kernel functions. SVM uses kernel methods to deal with nonlinear data because these methods create nonlinear combinations of the original properties of the data and make a mapping that separates these methods linearly and thanks to this mapping, a linear support vector machine algorithm transforms the training data into a higher dimensional feature area makes it trainable with. While performing all these operations, it uses kernel tricks since calculation cost will increase in cases where the size of the data set is too high.

C. Decision Tree Algorithm (DTA)

DT algorithms are used as a solution to both classification, and regression problems. It is a tree structure consisting of decision trees, decision nodes, and leaf nodes. Using the decision algorithm, it is started from the tree root, and the data on the most informative features are divided and the process is repeated until the node takes its purest form. Decision trees can be a tree with many nodes with this method, resulting in over-sleeping. In order to prevent overfitting in decision tree algorithms, a limit should be determined for the maximum number of nodes of the tree to be formed, and pruning should be done. The decision tree algorithm can be used to process both numerical, and categorical data, and it can process high-dimensional data in a short time due to its low computational cost. For this reason, it is preferred when data size increases among alternative algorithms.

D. K-Nearest Neighbor Algorithm (K-NN)

K-NN algorithm is an algorithm known as non-parametric sample-based learning. The model in this algorithm memorizes the training data set and finds the nearest neighbors as many as k number given to the algorithm in the data set, depending on the distance metric selected, and determines the classification tag of the data according to the majority. The K-Nearest neighbor algorithm allows easy adaptation for new data. However, it should be kept in mind that the cost of calculating these new samples increases in the worst case in direct proportion to the number of samples in the training data set, which is one of the main disadvantages of this algorithm.

E. Linear Discrimination Analysis Algorithm (LDA)

LDA is an unsupervised algorithm used to reduce the number of dimensions in the data set. LDA algorithms can also be used as a feature extraction technique to increase computational efficiency, and reduce overfitting due to "curse of dimension" in dirty data [19]. LDA computes the N-dimensional mean vector for each class by standardizing an N-dimensional data set creates the distribution matrix between and within classes. Converts matrices created here into low dimensional feature spaces. It is an algorithm that uses the covariance matrix to parse the data set. It works according to Fisher's score function, and estimates the linear coefficients that maximize the score. It tries to find linear combinations of variables that best describe the data and tries to model the difference between data classes. It can be used as a pre-process to extract features and identify classes prior to classification operations.

F. Gaussian Naïve Bayes Algorithm (GNB)

GNB algorithm is one of the Naive Bayes algorithm types. GNB is the type of supervised learning algorithm used when it has properties consisting of continuous values. In this algorithm developed on the basis of Bayes' theorem, each feature is considered independent from each other, which allows it to perform better than dependent models such as logistic regression. Good results can be obtained in both small size data, and high size data. Before using the Naive Bayes algorithm, the preprocessing and feature selection of the data should be done very well. Because there are not many parameters that can be optimized to change the result in the Naive Bayes algorithm, this optimization should be done first.

G. Logistic Regression Algorithm (LR)

LR algorithm is a supervised machine learning algorithm that is very easy to implement. It is used in regression problems as well as classification problems. It performs better in a classification that can be separated linearly. Logistic regression, which is a simple but powerful algorithm for linear and binary classification problems, gives better results especially when the variable type to be predicted is category type. The logistic regression algorithm establishes the relationship between the probability of a variable and its attributes in the data set. It deals not only with the class to be predicted but also with predicting the probability of the variable in the class. The algorithm uses the logit function to estimate the logarithmic rate of an event. The logit function takes input values between 0 and 1 and converts these data to values across the entire range of real numbers to represent a relationship.

H. Random Forest Algorithm (RF)

RF is an algorithm used in both classification and regression problems that can prevent overfitting by creating more than one decision tree. Regardless of whether the data is separated linearly in the random forest algorithm, the most important

features are inferred with all the decision trees created. Unlike the classical decision tree algorithm, the random forest algorithm tries to find the most important one in a random feature subset instead of finding the most important feature when dividing a node into sub-nodes.

VI. Results

Before performing image transformation over time series with the sonar signals in the data set, classical classification algorithms such as; DT, K-NN, LDA, Gaussian NB, LR, RF, and SVM algorithms have been trained and tested for this problem. Each algorithm grid-based hyperparameter search method was applied, and results were noted for the reserved and unseen test set. In order to identify the features for the classical machine learning algorithms, energy levels obtained from 60 frequency bands were used as features in the data set. For classical machine algorithms, feature selection or feature engineering methods were not performed. A new dataset is created with the pictures obtained with image transformation over time series methods in the same order as the case of the classical algorithms. In both cases, 75% of this dataset was used for training and 25% for testing. All the algorithms' performance was measured with accuracy, precision, recall and F_1 – Score metrics and given in Equations (11-14).

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \tag{11}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \tag{12}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP+TN}} \tag{13}$$

$$F_1 - \text{Score} = \frac{2\ Precision+Recall}{Precision+Recall} \tag{14}$$

In these equations, T_P represents the number of the actual positive values, T_N, the true negative values, F_P the false positive value and the F_N the negative value. Using these values, precision, recall, accuracy and F_1 criterion values are calculated for each algorithm to measure their performances.

The radial based function and C = 1.5 parameter values, which are among the nonlinear SVM algorithm kernel methods, have been studied, and the highest success rate has been achieved. DT algorithm parameters used in this study and giving the best results; Gini Information Criterion, the minimum number of instances required to split a node is 2, the minimum number of instances in a leaf node is 1, the maximum leaf node is not used, the class weight is not used, and the complexity parameter is 0.0. K-NN algorithm achieved the highest success rate when the value of k was trained as 1 with the data set in this work. Uniform value was given as weight, and auto was chosen as the algorithm, enabling it to decide the most appropriate algorithm according to the values. Since the data set in this study has many features, while working with the parameters of the LDA algorithm, the covariance matrix was not calculated using singular value decomposition (SVD) as a solvent. In addition to this parameter, the tolerance value was selected as 1.0e-4 because the other parameters were used as shrinkage auto, priority none, the number of components for dimension reduction none, covariance storage false and the solvent SVD. Random forest, parameters to achieve the best results with the data set in this study, the division criterion entropy, the maximum depth 15, the logarithm of the function required to find the maximum number of features, the minimum number of samples in a leaf node is 5, the minimum number of samples required to divide a node is 3, The number of trees to be created was 1000, and the number of processors was unlimited. All the results obtained from these algorithms are given in Table 2. Among the success rates, the highest result with 0.90% was SVM and the random forest algorithms with 0.86%.

Table 1: Comparison of results obtained from studies in literature review

[10]	Nearest Neighbor Classification – 12	90.40
[11]	[illegible]	[illegible]
[12]	RS1	49.50
	RS2	51.95
	RSU	51.40
	TS1	51.93
	TS2	53.43
	TSU	49.98
	RW1	51.93
	RW2	52.88
	RWU	51.38
[13]	C4.5 with IG	81.0
[14]	NBup with iDCM-CA	82.50
	SVM with iDCM-CA	60.0
	NN with iDCM-CA	88.75
	DT with iDCM-CA	96.25
	IBK with iDCM-CA	85.0
	LWL with iDCM-CA	90.25
[15]	[illegible]	[illegible]
[16]	[illegible]	[illegible]

Gaussian NB with the data set used in this work, the previous probabilities of the classes were not specified in the parameters studied in order to achieve the most ideal results, and the smoothing value required for calculation stability was chosen as 1.0e-9. Logistic regression, in this study, in order to achieve the most ideal results with the data set, the parameters were chosen as penalty norm L2, the solver as "linear" because the size of the data set was not high, the maximum iteration number was 100, and the warm start value was chosen not to be used because the solver value was linear.

Among the deep learning techniques, a convolutional neural networks architecture, which is quite successful and popular in image processing, was established, trained, and tested. The proposed custom convolutional neural network architecture is trained separately for each transformation process after data transformations are completed. Binary cross entropy loss function with L_2 regularization is used for optimization and for training, Adam optimization algorithm with learning rate 1.0e-5 was chosen and iterated over 100 epochs. Throughout

the iteration, the learning rate has been decreased to a certain extent and it has been tried to reach the minimum value in the loss function.

Table 2: Classical machine learning algorithms results of sonar signals in the data set with classification algorithms

SVM		0.90			
Random Fore		0.86			
Logistic		0.83			
Gaussian Naïv		0.67			
Decision Tre		0.71			
K-NN		0.86			
				0.80	0.84
LDA		0.76		0.77	0.77
				0.75	0.75

The success rates obtained with this convolutional neural network trained after RP, MTF and GADF transformations were ***95.3%, 96.65% and 95.43%,*** respectively. All the results and metrics obtained with this proposed method are given in Table-3.

Table 3: Time series – Image transformation method and the best results of the proposed custom CNN architecture

	Accuracy	Precision	Recall	F_1
RP	0.9434	0.9340	0.9620	0.9477
MTF	0.9323	0.9565	0.9212	0.9385
GADF	0.9057	0.9167	0.8798	0.8979

VII. Conclusion

This study proposed a novel information representation method for time series to obtain higher performance for classification problems without using any feature engineering and feature selection methods. The proposed method consists of converting regular time series to an image representation by implementing MTF, RP, and GAF transformations. After these transformations, spatial and temporal properties of time series are preserved and information stored as regular images.

By comparing the performance of classical ML and DL based time series-image transformation methods, it was seen that the proposed method could reach the best accuracy, precision, recall, and F1 score values of all. From Table 2 and Table 3, it was determined that the counting based probabilistic image transforms give better results for this data set. It is thought that the proposed method can be used as a more efficient feature extractor than classical methods thanks to RP, GAF, and MTF transformations, and it is the transformation method that takes into account the spatial correlation. These results prove that usage of the computer vision based DL algorithms without doing any feature engineering or selection can be proper for classification problems, especially for time series data.

References

[1] A D Waite *Sonar for Practising Engineers* Wiley Chichester; 2002

[2] L Jing *The principle of side scan sonar and its application in the detection of suspended submarine pipeline treatment*, Materials Science and Engineering, {IOP}, 439, 2018

[3] V L Lucieer Object-oriented classification of sidescan sonar data for mapping benthic marine habitats International Journal of Remote Sensing 29(3), 905–921, 2018

[4] A Burguera, G Oliver High-Resolution Underwater Mapping Using Side-Scan Sonar, PLOS ONE 11(1), 2016

[5] H J Flowers, J E Hightower A Novel Approach to Surveying Sturgeon Using Side-Scan Sonar and Occupancy Modeling Marine and Coastal Fisheries 5(1), 211–223, 2013 Available from: 10 1080/19425120 2013 816396

[6] A T Çelebi, M K Güllü, S Ertürk Mine detection in side scan sonar images using Markov Random Fields with brightness compensation 2011 IEEE *19th Signal Processing and Communications Applications Conference* (SIU), 916–919, 2011

[7] C S Huebner Evaluation of side-scan sonar performance for the detection of naval mines, Target and Background Signatures IV, Vol 10794 SPIE, 158–166, 2018 Available from: 10 1117/12 2325642

[8] M Verleysen, D François The Curse of Dimensionality in Data Mining and Time Series Prediction Lecture Notes in Computer Science, 3512(06), 758–770, 2005 Available from: 10 1007/11494669_93

[9] R P Gorman, T J Sejnowski Analysis of Hidden Units in a Layered Network Trained to Classify Sonar Targets Neural Networks 1(1), 75–89, 1988

[10] B Erkmen, T Yıldırım Improving classification performance of sonar targets by applying general regression neural network with PCA *Expert Systems with Applications*, 35(1-2), 472–475, 2008 Available from: doi:10 1016/j eswa 2007 07 021

[11] M Hossin, F Mahudin, I Din, A R Mat Analysis of Nine Instance-Based Genetic Algorithm Classifiers Using Small Datasets Journal of Telecommunication, Electronic and Computer Engineering, 9, 3–11, 2017 Novakovic J Using Information Gain Attribute Evaluation to Classify Sonar Targets 2009

[12] S Fong, S Deb, R Wong, Underwater Sonar Signals Recognition by Incremental Data Stream Mining with Conflict Analysis International Journal of Distributed Sensor Networks 10(5), 2014 Available from: 10 1155/2014/635834

[13] X Hong, J Zhang, S U Guan et al Incremental Maximum Gaussian Mixture Partition for Classification In: 2nd Joint International Information Technology, Mechanical and Electronic Engineering Conference (JIMEC 2017); Vol 62; March; 2017

[14] T Shang, X Xia, J Zheng MIME-KNN: Improve KNN Classifier Performance Include Classification Accuracy and Time Consumption DEStech Transactions on Computer Science and Engineering 2018; Available from: 10 12783/dtcse/csse2018/24490

[15] Z Wang, T Oates Imaging Time-Series to Improve Classification and Imputation, Proceedings of the *Twenty-Fourth International Joint Conference on Artificial Intelligence* (IJCAI), July 2015

[16] Z Wang, T Oates et al Spatially Encoding Temporal Correlations to Classify Temporal Data Using Convolutional Neural Networks Journal of Computer and Systems Sciences 2015

[17] J P Eckmann, S O Kamphorst, D Ruelle Recurrence Plots of Dynamical Systems Euro- Physics Letters (EPL) 4(9), 973–977, 1987

[18] S Raschka, V Mirjalili *Python Machine Learning*, 2nd Edition 2017 September Available from : https://sebastianraschka com/books html

A Research on Deep Learning Models Used in Turkish Text Analysis

G.KAYABAŞI KORU[1] and Ç.ULUYOL[2]

[1] Gazi University, Ankara/Turkey, glsmkayabasi@gmail.com
[2]Gazi University, Ankara/Turkey, celebiulu@gmail.com

***Abstract* – Deep learning approach in the fields of machine learning and artificial intelligence has facilitated the solution of many problems and has started to be used frequently. In today's technology, sharing opinions on various topics on social networking platforms, comments made on products on web blogs or e-commerce sites have made the issue of analysis of Turkish texts even more important. The transfer of users' thoughts and experiences through these platforms has helped companies to have an idea about their consumers. For this reason, the classification of short texts such as sentiment analysis, product reviews, movie reviews and tweets has become a focus. In this study, the deep learning approach used for Turkish text analysis will be mentioned and the models with effective results using this approach will be compared. The study is a survey article.**

***Keywords* – deep learning, natural language processing, turkish text analysis, machine learning, artificial neural networks,**

I. INTRODUCTION

DEEP learning is an important current subfield of artificial intelligence and machine learning topics that are very popular today. Deep learning methods are generally developed on artificial neural network studies. However, unlike these studies, they are based on more hidden neurons and layers [1]. Deep learning methods; It has yielded very successful results in processing many data types such as video, audio and text [2]. While mostly one method can be successful in text data processing such as natural language processing, a different method can give more successful results in processing video and audio data [3]. In some studies, suggested deep learning approaches can be successfully used in multimodal learning (including different modes such as text, image and sound at the same time) [4]. Among the application areas of deep learning methods; Language modeling and natural language processing, speech and speech processing, information retrieval, object recognition and computer vision, as well as multi-modal and multi-tasking learning [5].

Natural Language Processing (NLP) is a way of analyzing and obtaining meaningful information between the languages people speak and the computer language. Although it is generally regarded as a sub-science of computer science and linguistics, deep learning approaches, which are frequently used in this field recently, largely exclude the necessity of linguistics. The purpose of NLP is to examine problems in automatically creating and understanding natural languages [6]. NLP helps to ensure human-computer interaction by processing voices and texts produced by humans.

NLP are theoretically motivated computational techniques for automatic analysis and representation of human language. In an environment where people of all ages can access social media, the amount of data produced continues to increase day by day. Data created naturally by humans are not in a state to be directly processed. Therefore, the effort to make sense of data and use it efficiently in order to provide human-machine communication required it to work with many areas. While NLP studies in the early days focused on individual words, now they have focused on the relationship of words with each other and semantics studies on the whole [7].

NLP studies first go through a preprocessing that extracts features from sound or text to understand the text. Then, morphology, syntax, semantics and speech processing studies can be carried out. These fields of study have some difficulties in terms of word roots, word contexts and semantics. Grammar-rule-based NLP studies developed to overcome difficulties are based on handmade features [8]. Handmade features take time and are insufficient [9]. Deep learning, which is an important application of artificial intelligence methods developed to overcome the deficiencies, has provided better results with its artificial neural network structure, powerful hardware and big data input [10].

In this study, the deep learning methods used to do NLP will be explained, the third part will talk about the natural language processing problems performed

on the text, the fourth part will mention the studies made for Turkish texts and finally, the evaluations will be made about them.

II. DEEP LEARNING METHODS

This section discusses several popular deep learning networks. Table 1 contains a summary of the deep learning networks introduced in this section, their key key points and the most representative documents.

Table 1: Brief summary of deep learning networks

	Specifications
Deep Neural Networks	It is applied in many areas, but it can slow down the learning process with its application area.
Deep Autoencoders	It is a special artificial neural network used for unsupervised learning [11].
Deep Belief Networks	It is a specialized class of Boltzman machine that can operate with supervised and unsupervised.
Deep Boltzman Machine	It can be used with supervised and unsupervised, but the training process can be slow.
Generative Adversarial Networks	Unlike other machine learning algorithms, it has the ability to generate as well as learn.
Variational Autoencoders	It is an unsupervised learning based probabilistic graph model.
Recursive Neural Networks	These networks are used to model Sequential data.
Recurrent Neural Networks	It is good at identifying sequential patterns. Using LSTM (Long Short-Term Memory) for natural language processing yields successful results.
Convolutional Neural Networks	It is mostly used on computer vision and visual data. Even if big data is needed, it gives successful results. It is also good at natural language processing problems.

A. Deep Neural Networks

In today's conditions, the increase in data volume, the large number of variables in the data set, and efforts to analyze the complex and hierarchical relationships between variables have increased the interest in artificial neural networks-based deep learning approach. Deep neural networks are defined as the structurally deepened and expanded form of artificial neural networks. In other words, deep neural networks are the cases where the number of hidden layers in artificial neural networks is more than two and the number of neurons in these layers is high [12]. This network has an input layer and is directly connected to the output. With this algorithm, classifiers can be developed for the separation of patterns that can be separated linearly [13].

B. Deep Autoencoders

Autoencoders are one of the most popular models in Deep Learning area. As the name suggests, it aims to learn to convert any data to code automatically. It consists of 2 parts: Encoder and Decoder. These 2 parts are trained together as if they were a single model during the training phase. After the training, these models are separated and used to compress and decompress the data. For example, if the data is to be moved from one place to another, less data can be carried by installing an Encoder on the sender and a Decoder on the receiver. The purpose of the autoencoder model is to compress the given data and regenerate it with as little loss as possible. Therefore, the loss function of the system is the difference between output and input. The generated code is obtained by taking the output of the middle layer. This layer is called the bottleneck layer and determines the size of the code to be obtained. In other words, this layer determines how much compression will be made.

C. Deep Belief Networks

Deep belief networks (DBN), proposed by Geoffrey Hinton et al., Are created by stacks of Restricted Boltzmann Machines. In other words, the object in the stacks created by RBM is used in fields such as video sequences and motion capture. In short, it is a neural network approach that consists of successively added RBM layers. It is realized by training RMBs respectively. Deep belief networks architecture consists of many RBMs. DBNs are divided into more than one sub-branch and the usage

purposes of each sub-branch are processed in different ways. Therefore, these studies have been included in many different ways in many different fields. Studies are carried out on this subject under the following headings: Image Identification, Video Sequences, Motion Capture, Speech Recognition.

D. Deep Boltman Machine

A Boltzmann machine is a stochastic recurrent neural network type. Boltzmann machines are seen as the stochastic, productive counterpart of Hopfield networks. These are among the first neural networks that can learn internal representations and can represent and solve difficult combinatorial problems. Difference from DBN; deep boltzman machines have non-directional connections between all layers of the network. Another difference is that the time complexity is higher in deep Boltzmann machines and therefore the learning process is slow in large data sets [13].

E. Generative Adversarial Networks

Adversarial generative networks (GAN) were first introduced at the NIPS conference in 2014 by Ian Goodfellow and his working group. Adversarial generative networks are not traditionally supervised learning algorithms, although there have been some examples in the field of supervised learning lately, adversarial generative networks can be classified as unsupervised learning algorithms in terms of definition and behavior. To elaborate a little more, adversarial generative networks consist of two different network structures. The first of these is the generative network and the other is the discriminating network. While the generative network creates new data from a sample data set, the discriminating network tries to distinguish real and fabricated samples by looking at the data generated with the source data set.

F. Variational Autoencoders

New data can be generated by using Variational Autoencoder (VAE), which is a type of Autoencoder, as is done in GANs. Unlike Autoencoders, VAEs learn a certain probabilistic distribution on the bottleneck layer. In this way, new data can be generated from this distribution by using arbitrary codes. For example, by giving cat photos to the model, new cat photos or 3D models can be produced new 3D models and their resolutions can be increased.

G. Recursive Neural Networks

Similar to RNNs, recursive neural networks are natural mechanisms for modeling sequential data. This is because language can be viewed as an iterative structure in which words and subexpressions form other top-level expressions in a hierarchy. In this structure, a non-terminal node is represented by the representation of all its sub-nodes. The basic recursive neural network combines components in a bottom-up approach to calculate the representation of higher-order expressions. In one variant, MV-RNN (Multi-View Recurrent Neural Network), words are represented by both a matrix and a vector, meaning that the parameters learned by the network represent matrices of each component (word or sentence). Another variation, the recursive neural tensor network (RNTN) allows more interactions between input vectors to avoid large parameters as in the MV-RNN. Recursive neural networks show flexibility and are combined with LSTM units to deal with problems such as gradient fading. Recursive neural networks are used for a variety of applications: decomposition, using phrase-level representations for sentiment analysis, classification of semantic relationships (eg topic-message), sentence relationship.

H. Recurrent Neural Networks

Recurrent neural networks (RNNs) are networks in which connections between units form a directed loop. With RNN, dynamic temporal behavior is allowed. RNNs can use their input memory to manipulate inputs. This feature makes RNNs a usable method in handwriting recognition and speech recognition. People don't learn new meanings for a word they just learned. They attribute meaning to the newly learned word based on similar previously existing words. However, traditional artificial neural networks do not have this signification feature found in humans, and this is their biggest shortcoming. For example, when you want to classify the activities by looking at all the frames in the video, traditional neural networks will not be able to classify them as they cannot make sense between frames like humans. RNN, on the other hand, will create a loop, enable the use of past information, and thus make classification by making sense between frames. In traditional artificial neural networks, results from cells do not come back to

them as inputs. In RNN, the result coming out of the cell comes back to itself as an input. In the period of time, the same cell repeats itself more than once. Thus, interpretation can be established between the frames. Since RNNs can form a cycle, they can make sense of sequentially developing activities with each other. It has been widely used in recent years due to the ability to classify activities in the flow. RNNs have many uses; speech definition, language modeling, language translation, image captioning and so on.

İ. Convolutional Neural Networks

Another type of deep learning is convolutional neural networks (CNN), in which each module consists of a convolutional layer and a pool layer. These modules are usually stacked on top of or on top of a deep neural network to create a deep model [14]. The convolution layer shares many weights, and the pool layer downsamples the output of the convolution layer and reduces the data rate from the layer below. Weight sharing in the convolutional layer, together with properly selected pooling schemes, gives the CNN some "invariance" properties. It has been suggested that such limited "invariance" or equal variance is not sufficient for complex pattern recognition tasks and more principled ways of dealing with a wider range of invariance may be needed [15]. More recently, CNN has been found to be effective for speech recognition, with appropriate modifications from the CNN designed for image analysis, taking into account speech-specific features [16].

III. NLP PROBLEMS ON TEXTS

Among the deep learning methods, especially recurrent neural networks and convolutional neural networks are very successful in solving natural language processing problems on text. Including these two methods; It has been mentioned in different studies in the literature that other deep learning methods can be used in various problems of natural language processing on text and successful systems can be developed through these methods.

Deep learning methods used to solve natural language processing problems on the texts are shown in Table-2.

Table 2: Deep learning methods solving NLP problems on texts

NLP Problems on Texts	**Used Deep Learning Methods**
Text Classification	-Convolutional Neural Networks -Recurrent and Convolutional Neural Networks -LSTM ve Convolutional Neural Networks
Text Parsing	- Convolutional Neural Networks
Information Extraction	-Deep Neural Networks
Entity name Recognition	-LSTM ve Convolutional Neural Networks
Incident Extractiion	- Convolutional Neural Networks
Word type labelling	-Deep Neural Networks -LSTM

A. Text Classification

In order to automatically predict the author, subject, gender of the author and the mood of the author, the so-called automatic text classification has been carried out. In these studies, it is predicted to which class a new text belongs by examining the various properties of the texts with a certain class. However, in order to do this, it is necessary to decide which features of the texts will be examined [20].

There are successful studies on the text classification problem in the literature. For example; author recognition [21,22], text subject determination [23,24], e-mail classification [25], gender determination [21] can be given. In these studies, various representation methods are proposed for each problem type. The most used methods; word / phrase frequencies [23], ngram frequencies [25,26], word clustering [23], hidden meaning indexing [27], and functional words [28]. In another study, many methods in the literature and the methods proposed by them (multidimensional scaling and semantic space [29], word clustering based on class knowledge) were compared on various data sets [20].

B. Text Parsing

Text parsing is an area where natural language processing has been studied for many years and has many different methods in the literature to solve it. It is aimed to reveal the grammatical structure of a given text in text parsing [18]. In the literature, a fast method for text decomposition based on convolutional neural networks has been proposed. Here, convolutional neural networks are combined with structured label extraction in graphs, and thus

the graphical transformer network method, previously available in the literature, is used. The proposed method is open to improvement, but it is an effective method for text parsing [19].

C. Information Extraction

Another important area of research of natural language processing in text is information extraction. From the texts; Many important pieces of information such as assets, events, date and time expressions, concepts are extracted automatically. In the literature, a general approach independent of the type of information is proposed and in this study, deep neural networks are used for information extraction. In the study, it is provided to learn the features instead of creating the features manually. The study has been tested in different data sets and successful results have been obtained [30].

D. Entity Name Recognition

A named entity is a word or phrase that clearly defines an item from a set of items with similar properties [31]. Examples of named entities are generally institution, person, and location names; such as gene, protein, drug and disease names in the biomedical field. Entity name recognition is the process of finding and classifying named entities in the text according to predefined entity categories [32]. In one study, a hybrid and bi-directional approach for entity name recognition is proposed. For this, LSTM, an algorithm of recurrent neural networks, and convolutional neural networks are used. This hybrid method used determines the word level and character level features automatically [33].

E. Incident Extraction

It is desired to extract some information from the text for various purposes. Special mathematical studies have been carried out to reveal the events in a given text, verb structures in Turkish, word groups indicating the event, tense patterns and possible relationships between all these word groups [34]. The inference of the events in the texts is the subfield of information extraction. There are studies using convolutional neural networks for extracting sentence level features. The word-level features used were extracted using different word representations [35].

F. Word Type Labelling

It can be classified into types of words such as verbs, prepositions, and pronouns. Representations are made at the word and character level and convolutional neural networks are used for feature extraction. It has given effective results on various languages [36]. There are also language-independent studies in this area. Using LSTM networks, success has been achieved on different languages [37].

IV. STUDIES RELATED TO TURKISH TEXT ANALYSIS

Natural language processing has recently become popular with the increase in text data in line with the big data concept. Natural language processing techniques have been used extensively in the operations on the text [44]. Today, most of the work in natural language processing has been developed in the English language. The number of studies on natural language processing on the Turkish language is limited. In this section, studies in the literature for the Turkish language will be discussed. For Turkish text classification, a study has been conducted to compare the performance of representation methods on 6 Turkish data sets and prove that the n-gram method is the most successful among the word representation methods used in their studies [45].

In a study, the application developed in order to find spelling errors in Turkish texts was mentioned. In this study, spelling errors in Turkish texts were checked, foreign words in Turkish texts were found and Turkish equivalents were suggested for the found foreign words, misspelled words were identified and the found spelling errors were corrected [38]. In the literature, a Turkish data set consisting of 6 categories was created by collecting data from various news sites named TTC-3600. In addition, they made text classification with different machine learning algorithms on this data set [39]. The same dataset has been used to classify Turkish news texts by using deep learning and Word2Vec pre-processing methods, which have been studied on methods that are not encountered much in the literature. For this purpose, the TTC-3600 data set was used and classified with Convolutional Neural Networks [40]. Two deep learning-based models are designed for classification. The data set used was pre-processed before classification in these models

[43]. After the pre-processing step, the data set was enriched by using Word2Vec, one of the word vector representation methods, to increase the classification accuracy [41,42]. There is a study that analyzes emotions according to the content of Turkish Tweets. Words are represented by vectorization and classified by machine learning algorithms [46]. In another study, two data sets were created, the first data set was created for the training of word vectors, and the compilation was made from Turkish articles written in the Wikipedia environment with the general topic. In the other data set, news texts in 4 categories that are available on the internet in various newspapers and news sites were created. These data sets were vectorized using the word2vec method and the vector values were transferred to the numerical weights of the words in the data set to be classified by transfer learning. The data transformed from word to numerical values are classified by training with deep neural networks [47].

A new method was proposed in a study that was classified according to text emotions. In the approach called Supervised Meaning Classifier(SMC), very little training data such as 1% were used and the success of different approaches was calculated. In the study where classification methods such as Multinomial Naive Bayes (MNB) and SVM were compared, it was seen that the SMC method was more successful. While the training data set was at 70%, the SMC method showed 84% success, while SVM achieved 69% success. In the performance analysis of very low training data, which is the main subject of the study, while the training set was at 1%, SMC achieved 48% success, SVM achieved 19% success, and the method suggested by the authors was observed to be more successful [48].

In a study using text mining methods in Turkish documents, SVM, Sequential Minimal Optimization (SMO) and unigram performed bigram performance analysis [49]. In another similar study, they categorized newspaper writers according to their style and gender using n-gram techniques and compared their performance between n-grams at character level (character level n-gram) and n-grams at word level (Word level n-gram). According to their results, it was seen that SVM performed better than Naive Bayes and Random Forest approaches. They classified their column articles into the categories of economy, media and sports with an f-criterion score of 0.68 [50].

In another study, sentiment analysis from tweets was made using PV-DBOW and PV_DM models. It has been compared in Turkish and English documents. Emotions are classified as positive, negative, and neutral. Data such as Twitter usernames, names, html tags of people were not included in the training set. As the number of sentences increased, success rate measurements were made. According to the results obtained, it was seen that the PV-DBOW method is more successful than the PV_DM method [51]. A study has been conducted to classify whether there is spam in incoming SMS by using word2vec method. Random forest algorithm was used together with Word2vec method and they classified SMS successfully [52]. In another study, they classified the music genres and compared doc2vec and wrod2vec methods with classical BOW and CBOW approaches. Classical methods have been observed to be more successful according to the results obtained [53]. In his study, in which irony was determined with Turkish text, types of irony such as verbal irony, sensory irony and dramatic irony were determined, and ironies were tried to be determined with punctuation marks. Decision tree, Naïve Bayes, Logistic Regression, Decision Table, Random Forest, Multilayer Perception approaches were compared and it was determined that the most successful algorithms in detecting irony were Random Forest, Multilayer Perception [54]. In this study [55] comparing the performances of plagiarism detection software, it was determined that the presence of Turkish characters in the texts reduces the performance of plagiarism detection.

In the study where the texts were classified in the fields of economy, magazine, health, politics and sports, [56] a classification was made based on the frequency of occurrence of words in a text. In the classification, the performances of Naïve Bayes (NB), Support Vector Machine (SVM), KNN and Random Forest (RF) methods were compared. The highest success was achieved in Naïve Bayes with 96%. While the most successful results in classification are achieved in the field of economics, it has been determined that the most wrong area is politics. It has been observed that some words in the field of health and politics also appear in the field of

economics, which reduces the performance of the classification.

In another study conducted with the columnists of the authors selected from the Sabah newspaper, the profile of the authors was created and the study of which author their columns belonged to. In the study, the punctuation marks, spaces, and numbers of landing on the bottom line were used by the authors to distinguish them [57]. Differences have been identified, such as some authors rarely use exclamation marks, some authors use exclamation marks more frequently, and some use more commas because they prefer long sentences. Similarity analysis was performed between the vectors created with these data and the texts of the authors, and they achieved a success rate of 86%.

In a study conducted on Hürriyet newspaper upon the determination of the author, texts belonging to 18 different authors taken from the magazine health and politics were studied [58]. 20 columns were selected from each author and a performance analysis was made on these articles. In the approach put forward by the authors; The number of words in each text, number of sentences, average word length, sentence length, number of different words, question marks, semicolons, exclamation marks, number of punctuation marks, number of lines as well as grammatical features of sentences were determined as distinctive parameters. Accordingly, 22 different distinctive style parameters were extracted along with features such as the average number of nouns, average number of adjectives, and average predicates in the texts of each author. A success of 67% was achieved by comparing the authors on the vector created over these 22 features. They achieved a success of 60% with the Multilayer Perception (MLP) method, the approach put forward by the authors, and a success of 72% with the Radial Base Function approach. In the study, the authors listed the features that were most successful in separating them from each other as follows: the average number of words in a sentence, the average word length used, the average number of nouns, the average number of adjectives, the average number of predicates, the average number of conjunctions, the average number of pronouns, the number of points, the number of incomplete sentences. the ratio of all sentences, the ratio of inverted sentences to all sentences.

[59] In his study, it was tried to determine the subject of instant messages. In the study, in which the emojis used in speeches and their abbreviations were taken into consideration, the titles were determined as sports, love / marriage, education, entertainment, swearing and other. Naïve Bayes, k-NN and SVM classification methods were used to separate the texts. The most successful result was determined as swearing at 92%. The reason for this is that swearing speech can be differentiated from the others due to the words it contains. [60] In his study, profiles of columnists were created with the data obtained from the Hürriyet newspaper. In the study in which the Paragraph Vector (PV) method was used, the performances of the PV-DM (Paragraph Vector-Distributed Memory) and PV-DBOW(Paragraph Vector-Distributed Bag of Words) models, which are among the paragraph vector models, were compared. It was concluded that the PV-DM model was more successful in distinguishing the authors from each other. In addition, it was observed that model performances decreased with the increase in the number of authors who were tried to be separated from each other. In this study, authors' vectors were created in 2014 using the doc2Vec method, which is an extended version of word2Vec, and author similarity predictions were successfully made in real time. In our study, the subjects written by the authors, the data obtained from the newspapers and Ekşi Sözlük, and their models were created and classified successfully [61].

V. CONCLUSION AND RECOMMENDATIONS

In this study, studies that use Turkish texts and create data sets are mentioned. Each study used unique natural language processing methods and achieved success. We examined these successful studies and prepared a document that you can find all of them together in order to contribute to the literature. In the later stages of our work, we will develop our own application with these natural language processing methods and present it to you.

REFERENCES

[1] Ravì, D , Wong, C , Deligianni, F , Berthelot, M , Andreu-Perez, J , Lo, B , & Yang, G Z (2017) "Deep learning for health informatics" IEEE journal of Biomedical and Health Informatics 21(1), 4-21

[2] Lecun, Y , Bengio, Y , & Hinton, G (2015) "Deep learning" Nature 521(7553), 436

[3] Collobert, R , & Weston, J (2008) "A unified architecture for natural language processing: Deep neural networks with multitask learning" International Conference on Machine Learning (ICML), 160-167

[4] Ngiam, J , Khosla, A , Kim, M , Nam, J , Lee, H , & Ng, A Y (2011) "Multimodal deep learning" International Conference on Machine Learning, 689-696

[5] Deng, L , & Yu, D (2014) "Deep learning: methods and applications" Foundations and Trends in Signal Processing, 7(3–4), 197-387

[6] Young, T , Hazarika, D , Poria, S , & Cambria, E (2018) Recent Trends in Deep Learning Based Natural Language Processing IEEE Computational Intelligence Magazine, 55-75

[7] Cambria, E , & White, B (2014) Jumping NLP curves: A review of natural language processing research IEEE Computational Intelligence Magazine, 9(2), 48-57

[8] Bill, E (1992) A Simple Rule-Based Part of Speech Tagger ANLC '92 Proceedings of the third conference on Applied natural language processing Trento

[9] Young, T , Hazarika, D , Poria, S , & Cambria, E (2018) Recent Trends in Deep Learning Based Natural Language Processing IEEE Computational Intelligence Magazine, 55-75

[10] Song, H , & Lee, S -Y (2013) Hierarchical Representation Using NMF International Conference on Neural Information Processing 2013: Neural Information Processing, (s 466-473)

[11] Y Bengio, «Learning Deep Architectures for AI,» Found trends Mach Learn, cilt vol 2, no no 1, p 1–127, 2009

[12] Deep Neural Network https://www techopedia com/definition/32902/deep-neural- network

[13] Ravì, D , Wong, C , Deligianni, F , Berthelot, M , Andreu-Perez, J , Lo, B , & Yang, G Z (2017) "Deep learning for health informatics" IEEE journal of Biomedical and Health Informatics 21(1), 4-21

[14] Y LeCun, L Bottou, Y Bengio, and P Haffner Gradient-based learning applied to document recognition Proceedings of the IEEE, 86:2278–2324, 1998

[15] G Hinton A better way to learn features Communications of the Association for Computing Machinery (ACM), 54(10), October 2011

[16] T Sainath, A Mohamed, B Kingsbury, and B Ramabhadran Convolutional neural networks for LVCSR In Proceedings of International Conference on Acoustics Speech and Signal Processing (ICASSP) 2013

[17] Socher, R , Bengio, Y , & Manning, C D (2012) "Deep learning for NLP (without magic)" Tutorial Abstracts of ACL 2012, 5

[18] KÜÇÜK, D , & ARICI, N (2018) DOĞAL DİL İŞLEMEDE DERİN ÖĞRENME UYGULAMALARI ÜZERİNE BİR LİTERATÜR ÇALIŞMASI Uluslararası Yönetim Bilişim Sistemleri ve Bilgisayar Bilimleri Dergisi, 2(2), 76-86

[19] Collobert, R (2011) "Deep learning for efficient discriminative parsing" International Conference on Artificial Intelligence and Statistics, 224-232

[20] Amasyalı, M F , BALCI, S , Emrah, M E T E , & VARLI, E (2012) Türkçe metinlerin sınıflandırılmasında metin temsil yöntemlerinin performans karşılaştırılması/a comparison of text representation methods for turkish text classification EMO Bilimsel Dergi, 2(4)

[21] Amasyalı, M F , Diri, B , "Automatic Turkish Text 11th International Conference on Applications of Natural Categorization in Terms of Author, Genre and Gender", Language to Information Systems-NLDB2006, LNCS 11th International Conference on Applications of Natural Volume 3999, 2006

[22] Türkoğlu, F , Diri, B , Amasyalı, M F , "Author International Conference on Intelligent Computing, ICIC Attribution of Turkish Texts by Feature Mining", Third 2007, Qingdao, China, LNCS Volume 4681, 2007

[23] Bekkerman R , Ran El-Yaniv, Naftali T , Yoad W , Categorization", Journal of Machine Learning Research, "Distributional Word Clusters vs Words for Text 1-48, 2002

[24] Song, F , Liu, S , & Yang, J (2005) A comparative study on text representation schemes in text categorization Pattern analysis and applications, 8(1-2), 199-209

[25] Çıltık, A , & Güngör, T (2008) Time-efficient spam e-mail filtering using n-gram models Pattern Recognition Letters, 29(1), 19-33

[26] Ciya L , Shamim A , Paul D , "Feature Preparation in Text Categorization", Oracle Text Selected Papers and Presentations,2001

[27] Özel, B , "Küresel k-Ortalamalı Gruplama Yöntemi ile Metinlerin ve Terimlerin Saklı Anlam İndekslenmeleri", ASYU,Istanbul, Turkey, 223-227, 2004

[28] Holmes, D I , "The Evolution of Stylometry in Humanities Scholarship", Literary and Linguistic Computing, 13 (3): 111-117, 1998

[29] Amasyalı, M F , Davletov, F , Torayew, A, Çiftçi, Ü, "text2arff: Türkçe Metinler İçin Özellik Çıkarım Yazılımı", SİU, Diyarbakır,2010

[30] Qi, Y , Das, S G , Collobert, R , & Weston, J (2014) "Deep learning for character-based information extraction" European Conferenceon Information Retrieval, 668-674

[31] R Sharnagat, "Named entity recognition: A literature survey," Center For Indian Language Technology, 2014

[32] Li, J , Sun, A , Han, J & Li,C (2020) A survey on deep learning for named entity recognition IEEE transactions on Knowledge and DataEngineering

[33] Chiu, J P , & Nıchols, E (2015) "Named entity recognition with bidirectional LSTM-CNNs" arXiv preprint arXiv:1511 08308

[34] Sadi Evren SEKER, Banu DIRI, International Conference on Artificial Intelligence konferansı dahilinde, "Proceedings of International Conference on Artificial Intelligence", bildiri "TimeML and Turkish Temporal Logic", pp 881-887, ICAI 2010

[35] Chen, Y , Xu, L , Lıu, K , Zeng, D , & Zhao, J (2015) "Event extraction via dynamic multi-pooling convolutional neural networks" Annual Meeting of the Association for Computational Linguistics and International Joint Conference onNatural Language Processing, 167-176

[36] Dos Santos, C N , & Zadrozny, B (2014) "Learning character-level representations for part-of- speech tagging" International Conference on Machine Learning (ICML), 1818 1826

[37] Plank, B , Søgaard, A , & Goldberg, Y (2016) "Multilingual part-of-speech tagging with bidirectional long short-term memory models and auxiliary loss" arXiv preprint arXiv:1604 05529

[38] Delibas, A (2008) Doğal Dil İşleme İle Türkçe Yazım Hatalarının Denetlenmesi (Doctoral dissertation, Fen Bilimleri Enstitüsü)

[39] D Kılınç, A Özçift, F Bozyigit, P Yıldırım, F Yücalar, E Borandag, "TTC-3600: A new benchmark dataset for Turkish text categorization",Journal of Information Science, 43(2), 174–185, 2017

[40] H H Aghdam, E J Heravi, Guide to convolutional neural networks : a practical application to traffic-sign detection and classification,Springer, A B D , 2017

[41] Y Goldberg, O Levy, "word2vec Explained: deriving Mikolov et al 'snegative-sampling word-embedding method", arXiv:1402 3722, 2014

[42] T Mikolov, I Sutskever, K Chen, G Corrado, J Dean, "Distributed Representations of Words and Phrases and their Compositionality", 26th International Conference on Neural Information Processing Systems, 3111-3119, Nevada, A B D , 2013

[43] Çiğdem, A C I , & ÇIRAK, A (2019) Türkçe Haber Metinlerinin Konvolüsyonel Sinir Ağları ve Word2Vec Kullanılarak Sınıflandırılması Bilişim Teknolojileri Dergisi, 12(3), 219-228

[44] Çoban, Ö , & Özyer, G T (2018, September) Word2vec and Clustering based Twitter Sentiment Analysis In 2018 International Conference on Artificial Intelligence andData Processing (IDAP) (pp 1-5) IEEE

[45] Amasyalı M F , Çetin M , Akbulut C , "Metinlerin anlamsal uzaydaki temsil yöntemlerinin sınıflandırma performansına etkileri," Sigma, vol 5, pp 8–14, 2013

[46] Ayata D , Saraclar M , Ozgur, A (2017) Turkish tweet sentiment analysis with word embedding and machine learning 1-4 10 1109/SIU 2017 7960195

[47] AYDOĞAN, M , & KARCI, A Makine Öğrenmesi ve Transfer Öğrenme ile Türkçe Metin Sınıflandırma Turkish Text Classification with Machine Learning and Transfer Learning

[48] M C Ganiz, M Tutkan, and S Akyokuş, "A novel classifier based on meaning for text classification," in 2015 International Symposium on Innovationsin Intelligent SysTems and Applications (INISTA), 2015, pp 1–5

[49] I Mayda and M Yesiltepe, "N-gram based approach to recognize the twitter accounts of Turkish daily newspapers," 2017, pp 1–5

[50] A Deniz and H E Kiziloz, "Effects of various preprocessing techniques to Turkish text categorization using n-gram features," in 2017 International Conference on Computer Science and Engineering (UBMK), 2017, pp 655–660

[51] M Bilgin and I F Senturk, "Sentiment analysis on Twitter data with semisupervised Doc2Vec," 2017, pp 661–666

[52] O Karasoy and S Ballı, "Classification Turkish SMS with deep learning tool Word2Vec," in 2017 International Conference on Computer Science and Engineering (UBMK), 2017, pp 294–297

[53] Ö Çoban and I Karabey, "Music genre classification with word and document vectors," in 2017 25th Signal Processing and Communications Applications Conference (SIU), 2017, pp 1–4

[54] O Dülger, "Türkçe Metinlerde İroni Tespiti "

[55] M Kaya and S A Özel, "A Comparison of Text Similarity Detection Software for Turkish Documents and Investigating the Effects of Stemming and Turkish Character Usage," Çukurova Üniversitesi Mühendis -Mimar Fakültesi Derg , vol 29, no 2, pp 115–130, Dec 2014
[56] H K Yildiz, M Genctav, N Usta, B Diri, and M F Amasyali, "A New Feature Extraction Method for Text Classification," in 2007 IEEE 15th Signal Processing and Communications Applications, 2007, pp 1–4
[57] H Takçı and E Ekinci, "Character Level Authorship Attribution for Turkish Text Documents," TOJSAT, vol 2, no 3, pp 12–16, Sep 2012
[58] B Diri and M F Amasyalı, "Automatic author detection for turkish texts," in Artificial Neural Networks and Neural Information Processing (ICANN/ICONIP), 2003, pp 138–141
[59] Ö Özyurt and C Köse, "Chat mining: Automatically determination of chat conversations' topic in Turkish text based chat mediums," Expert Syst Appl , vol 37, no 12, pp 8705–8710, Dec 2010
[60] M Sarı and A M Özbayoğlu, "Classification of Turkish Documents Using Paragraph Vector," in 2018 International Artificial Intelligence and Data Processing Symposium (IDAP), 2018, pp 1–4
[61] Sarı, M (2018) Derin öğrenme yöntemleri kullanılarak Türkçe doküman sınıflandırma (Master's thesis, TOBB ETÜ Fen Bilimleri Enstitüsü)

Interpretation of Turkish Sign Language with Deep Learning and Mobile Application

Y.ANAGUN[1]

[1] Eskisehir Osmangazi University, Eskisehir/Turkey, yanagun@ogu.edu.tr

Abstract - **According to the data of the Federation of the Hearing Impaired, there are more than 4 million hearing-impaired people in Turkey. The most important problem faced by hearing and speech impaired people is not to be understood by people without hearing impairment. In line with this problem, the primary aim of this study is to eliminate the communication problems of our hearing-impaired citizens by combining Turkish sign language (SL) with deep learning technologies and to increase the awareness of hearing impaired individuals in our society. In this study, a SL interpretation program has been developed, which we anticipate that productivity will increase with the widespread use of mobile devices. In this context, existing SL movements will be created with a model trained with deep learning and the model will be implemented in a mobile program. Hence, it is possible to translate the desired movement into Turkish.**

Keywords - **Sign language, Turkish sign language, deep learning.**

I. INTRODUCTION

ACCORDING to the data of the World Health Organization (WHO), approximately 466 million people in the world have hearing impairments, and 34 million of these are children [1]. Federation of Hearing Impaired (FHI) in Turkey describes this number at close to 3 million [2]. The biggest problem that people with hearing impairments experience in society is not being understood by people without hearing impairments. For this reason, they want SL, which is their native language, to be defined as a constitutional right. For example, in the field of health, some hospitals trying to find solutions to end the communication problems experienced by the hearing impaired have taken action and started to provide SL training to their employees, including physicians and nurses. For this purpose, health care workers in Turkey receive training in SL by FHI. Hearing and speech impairments citizens face various difficulties in communication not only in the field of health but also in daily life. In this study, it is aimed to minimize the communication problems experienced by the hearing-impaired people by making use of today's technologies and the problem of "not being understood". In literature, there are some studies on sign languages of different countries. [3] proposed a study combining deep learning and Support Vector Machines (SVM) to realize the transformation of SL into emotional speech. A continuous SL recognition framework with deep neural networks was developed by [4] that copies videos of SL sentences directly into sequences of ordered gloss labels. Another study that interprets Convolutional Neural Networks (CNN)'s outputs in a Bayesian framework has been developed in [5]. The Hybrid CNN and Hidden-Markov-Models (HMM) model has combined the powerful discrimination abilities of CNNs with the sequence modeling capabilities of HMMs. Turkey has less work in the field of recognition of SL. Although the studies performed or suggested are new, they are immature and trial phase studies. In Turkey, there are less work in the field of recognition of SL. Although the studies have been developed and suggested are up to date, they are immature and experimental studies. In [6], a real-time American sign language recognition method proposed using desk and wearable computer based video. [7] developed a Chinese Sign Language (CSL) recognition system using the portable and cost-effective Leap Motion sensor and the kth-Nearest Neighbor (k-NN) technique. In recent years, the CNN has been widely used in many methods to recognize sign language. In addition, the recent literature studies indicates that deep neural networks based methods are successful and robust for sign recognition. A new model based on deep CNN used by deaf people presented in [8] to recognize Russian alphabet letters presented as static signs in Russian Sign language. Weakly supervised learning based approaches can be carried out in video domain. [9] demonstrated this by adding strong CNN-Long short-term memory (LSTM) models to each HMM stream following the hybrid approach, adding interim synchronization constraints between streams. In terms of the subject, this study aims to eliminate the communication problem experienced by hearing impaired people and to create a communication tool that translates sign language movements, which they define as their native language into daily speech. Recognition of hand sign language from video is a challenging area of researchers due to some undesired factors such as rapid hand movements, brightness and background complexity. Deep learning-based pipeline architecture for efficient automatic hand sign language recognition using Single Shot Detector (SSD), 2D Convolutional Neural Network (2DCNN), 3D Convolutional Neural Network (3DCNN), and Long Short-Term Memory (LSTM) from RGB input videos proposed in [10].

This study can be a source of ideas for different social responsibility projects. According to the successful results to be obtained as a result, it is planned to bring it to a next level that can be used by hearing-impaired people. The rest of this paper is arranged as follows. Details of the proposed method and Experimental (visual) results are given in Section II and Section III, respectively. Finally, we conclude this paper in Section IV.

II. PROPOSED METHOD

This study aims to reduce the communication problems of hearing-impaired citizens. Further, in terms of social responsibility, a unique contribution to the literature will be made by keeping these people in touch with the society and integrating them into the society.

Sign language is a very rich communication tool today. Creating a data set from a whole language is quite difficult and time-consuming problem. Therefore, frequently used words in daily life have been chosen. The most important characteristics of the chosen words are that they give nouns or actions and contain as little action as possible. The main reason for the movement restriction is that we want to keep the number of frames to be captured as low as possible. In training phase, 8-10 frames were captured and a data set was created over them to understand of movement. The meaning of word is extracted from the current sign language from a video sequence using 3D-CNN model, Optical flow, RestAPI and Android technologies as a whole.

Firstly, all video records in the data set are converted into the frames. The video frames have been recorded with the tag number of the class to which the relevant movement belongs, and the name of the video obtained and filed. Then, motion vectors are estimated on these video frames by optical flow. The optical flow result is shown in color in Figure 1. As a result of the optical flow, the motions on frame are colored by obtained motion vectors. Thus, the region of body movements can be determined within that frame. At the end of this process, the colored frames to be used for training are recorded in files with the names of the relevant video frames. Training model is shown in Figure 2.

Figure 1: Optical flow result

A server environment has been created in order to test the trained model and use it via a mobile application. The video is instantly recorded and sent to the server from a mobile application. Then, by using the model running on the server, the motion is estimated and the response is returned to the mobile application. Video are converted into the frames and optical flow estimation performed on these frames at server environment. The obtained data is evaluated by the model and the motion in the video is labeled with the most similar motion and expressed the amount of similarity as a percentage.

In the client, there is an interface structure, ServiceGenerator and FileChooser class structures. Connection processes are carried out with the retrofit library. The structure of the data to be uploaded to the server is kept in the interface structure. In

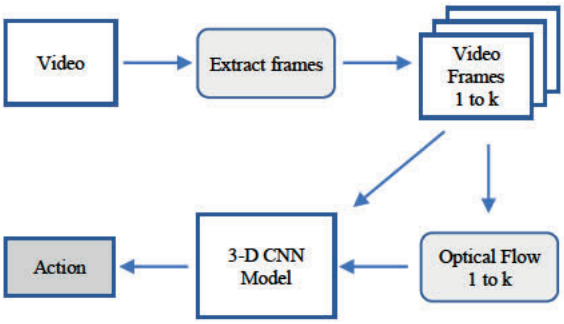


Figure 2: Training Model

ServiceGenerator class, there are URL information to be connected, TimeOut duration and connection settings. The target location information of the recorded video is determined in the FileChooser class. Once the target location of the video is specified, the content information is forwarded to the upload process. In the Main class, requests are sent to the server, camera access, permission operations and video sending are performed. Client-server relational diagram is presented in Figure 3.

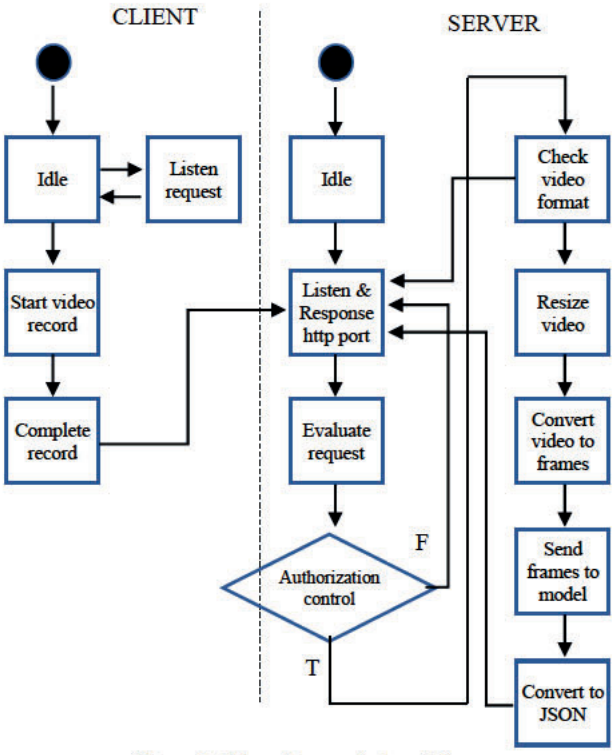


Figure 3: Client-Server relational diagram

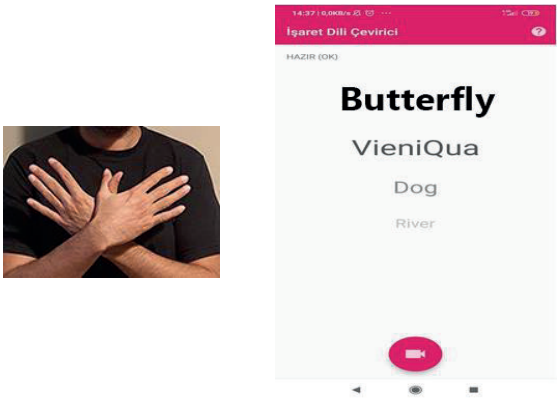
III. EXPERIMENTAL RESULTS

The CNN model is implemented by using a popular python library Keras with Tensorflow-backend. There are 20 layers in

total in the 3D-CNN model. All experiments are achieved on a GPU cluster set that consists of eight NVIDIA QUADRO P5000 16 GB memory. The modelling simulations are performed on seven NVIDIA QUADRO P5000 GPUs. The number of epoch is 17, batch size is 4 and ReLu is used as the activation function.

The data set used in model training is the Isolated Gesture Recognition (ICPR '16) data set created by the ChaLearn team of the University of Barcelona. The data set contains 47933 RGB-D motion videos [11]. These movements consist of 249 classes. This data set created does not consist of data belonging to a single sign language. It also contains data belonging to different sign languages.

The user interface is developed as an android based application as planned in the design phase of the study. Retrofit2 serves using okhttp3 and API 23 Camera API. The server has been created to meet client requests, running with the Flask library using Python 3.6 on a virtual machine with an Ubuntu operating system, behind the load balancer response offered by DigitalOcean. In Figure 4, possible motion data and result are shown in the Android-based user interface. There are the last 4 results presented in the user interface. Previously shown results are deleted from the screen.



(a) (b)

Figure 4: Experimental result (a) Butterfly test video (b) Mobile application result

IV. CONCLUSION

In recent years, deep learning technology has become an important technology for application development. This technology can be expressed as a kind of artificial intelligence that can detect features using large and numerous tagged training data. In this study, although the expected results are generally achieved successfully, various performance problems arise. First challenge, the application does not work efficiency in every environment. A key problem the optical flow motion vectors are miscalculated due to brightness. Another problem is that video resolutions can vary. Resolution is very effective in estimating the motion in videos. In low-resolution videos, the results cannot be generated accurately. In order to eliminate all these problems, various image processing techniques should be added to the study. For instance, in very light or very dark environments, such problems should be minimized and the success rate should be increased. Since there is no data set belonging to Turkish Sign Language in model trainings, signs were obtained from other sign languages in mixed form. For future works, it is planned to make sentence detection after determining possible words from all consecutive video frames.

REFERENCES

[1] https://www who int/news-room/fact-heets/detail/deafness-and-hearing-loss
[2] https://www ief org tr
[3] N Song, H Yang and P Zhi, “Towards Realizing Sign Language to Emotional Speech Conversion by Deep Learning”, In Pro of the Int Conf of Pioneering Comp Scientists, Engineers and Educators (ICPCSEE 2018), Zhengzhou, China, 21 - 23 September 2018
[4] R Cui, H Liu, and C Zhang, “A Deep Neural Framework for Continuous Sign Language Recognition by Iterative Training”, IEEE Transactions on Multimedia, vol 21, no 7, pp 1880 - 1891, July 2019
[5] O Koller, S Zargaran, H Ney, and R Bowden, “Deep sign: hybrid CNN-HMM for continuous sign language recognition”, Int J Comput Vis , 126, pp 1311 - 1325, 2018
[6] T Starner, J Weaver, A Pentland, “Real-time American sign language recognition using desk and wearable computer based video”, IEEE Trans on Pattern Analysis and Machine Intelligence, vol 20, no 12, pp 1371 - 1375, 1998
[7] Y Xue, S Gao, H Sun and W Qin, “A Chinese Sign Language Recognition System Using Leap Motion”, 2017 International Conference on Virtual Reality and Visualization (ICVRV), Zhengzhou, China, 21 - 22 Oct 2017
[8] I Makarov, N Veldyaykin, M Chertkov and A Pokoev, “Russian Sign Language Dactyl Recognition”, 2019 42nd International Conference on Telecommunications and Signal Processing (TSP), Budapest, Hungary, 1 - 3 July 2019
[9] O Koller, N C Camgoz, H Ney and R Bowden, “Russian Sign Language Dactyl Recognition”, IEEE Trans on Pattern Analysis and Machine Intelligence, vol 42 , no 9, pp 2306 - 2320, 2019
[10] R Rastgoo, K Kiani and S Escalera, “Hand sign language recognition using multi-view hand skeleton”, Expert Systems with Applications, vol 150, 113356, 2020
[11] http://chalearnlap cvc uab es/

Classification of Customer Complaints with Deep Learning

Y.ANAGUN[1]

[1] Eskisehir Osmangazi University, Eskisehir/Turkey, yanagun@ogu.edu.tr

Abstract **- The purpose of this study is to classify customer complaints of the financial sector with a deep learning model and to ensure that customer complaints are forwarded to the relevant institution or the relevant department of the institution in question. Nowadays, many companies perform manual classification procedures by company officials assigned on various platforms. Employees following these platforms read and evaluate complaints and forward them to the relevant department within the company. Since this complaint is a human factor in the classification process, the process is slow and troublesome. In this study, it is aimed to develop a web application that bank customers can share their complaints. The department where customer complaints are conveyed is determined by the deep learning model running on the server. This model are developed with Recurrent Neural Networks (RNN) architecture and the dataset includes Turkish complaints. Server communicates with a web application by RESTful API and then a complaint is sent to the relevant department.**

Keywords **- Turkish Text Classification, Recurrent Neural Network, Deep Learning.**

I. INTRODUCTION

WHEN a business receive complaints from their customers, the correct management of these complaints is as important as ensuring customer satisfaction. According to research on disgruntled customers, approximately 96% of these customers do not complain [1]. Many customers stop buying goods and services from that business or choose to get their requests from other companies. Therefore, considering that only 4% of discontent customers deliver their complaints to the business, it can be say that complaints received by the business about 25 times. In this way, the business should handle an incoming complaint as if there were 25 complaints on the same subject and plan the appropriate actions related to the complaint accordingly. In recent years with developing technology, businesses have easy access to mobile applications, web pages, customer services and other used complaint channels. However, customer complaints have been become varied and the number of customers reporting complaints have been increased. This case makes it difficult for businesses to return to customer complaints. Further, it causes businesses to lose customers and prevents them from gaining new customers. The purpose of this study is to classify customer complaints in the finance sector and ensure that customer complaints are delivered to the relevant department of the institution. In addition, by increasing the rate of return to complaints in the financial sector, reducing the time to return to customer as much as possible and increasing customer satisfaction.

In many companies, complaint classification is done manually by company officials assigned on various platforms. Employees follow these platforms read the complaint, evaluate it and forward it to the relevant department within the company [2]. The mentioned process brings some problems such as sending the complaint to the wrong department, not being evaluated correctly or customer complaint due to the human factor. If the classification process is desired to be done with the help of a computer, the key idea will be smart systems. There are various studies about text classification in the literature. Classification of news texts and sentiment analysis for Turkish can give example. The news texts classified and compared using the Turkish Text Classification 3600 (TTC-3600) dataset in [3]. According to this study, CNN and Word2Vec method showed better performance (93.3% accuracy) than classical statistical and machine learning based classification algorithms. Another approach that classifies texts using the Turkish data set and performs emotional analysis as positive, negative or neutral according to three classes is presented in [4]. They have developed a deep learning model that can achieve 95% accuracy using 243.000 customer reviews and GRU (Gated Recurrent Units) architecture. [5] trained and tested the CICIDS 2017 dataset with Distributed Denial of Service (DDoS) attack data by using Multilayer Sensors (MLP), LSTM (Long-Short Term Memory) and GRU networks. When the results of this study are examined, it can be understood that there is no significant difference between GRU and LSTM. Moreover, it is seen that LSTM and GRU, which are RNN (Recurrent Neural Network) types, give better results than MLP structure. [6] developed a hierarchical graph transformer-based approach for Large-Scale Multi-Label Text Classification. First, they modeled the text, a graphic structure that could embody the different meanings of the text and the connections between them. Then they use a multi-layer transformer structure with a multi-head attention mechanism at the word, sentence, and graph levels. [7] have developed a Tree-LSTM model, a generalization of LSTM to tree-structured network typologies, to learn rich semantic representations. [8] proposed a much simpler CNN-based model than Dynamic CNN (DNN) for text classification. [9] developed a new CNN-based model that makes two updates to the architecture of proposed in [8]. Character-level based CNNs have also been investigated for text classification [10, 11]. [12] developed a CNN architecture to classify text at character level that considerably reduced the memory consumption and required training time. There are several studies on literature examining the impact of word placing and CNN architectures on learning performance. [13] proposed a Very Deep CNN (VDCNN) model for text classification inspired by VGG [14] and ResNets [15]. [16]

demonstrated that deep models actually outperform shallow models when the text inputs are represented as a sequence of characters. In DenseNet [17], it has been shown that a simple shallow-and-wide network architecture performs better than deep models with word inputs. [18] worked on weighted word placing via a multi-channel CNN model. [19] presented a tree-based CNN to construct sentence level semantics. [20] used multi-layer CNNs to determine salient n-gram patterns and text matching patterns as the image recognition task. [21] developed a CNN-based model that combines implicit and explicit forms of short text for classification. More recently, methods using CNNs have found great success to classify the biomedical text [22-25].

II. PROPOSED METHOD

In this study, especially the financial sector customer complaints which are received through the web application interface are evaluated. If the customer complains is more relevant about which department of the company, it is delivered to that department with classified form without any human intervention. Classification of complaints is achieved by deep learning model. GRU and LSTM architectures are used in text classification. RNN architectures offer an approach based on previous knowledge usage. It is possible to say that the RNN architecture gives satisfactory results in Turkish news text classification and emotional analysis.

The proposed method consists of the operation of customers sharing the complaints of the banks using the web application, classifying the shared complaints with the deep learning model and communicating them to the relevant employee of the bank and responding to the complaint. The conceptual data flow diagram of proposed method is shown in Figure 1.

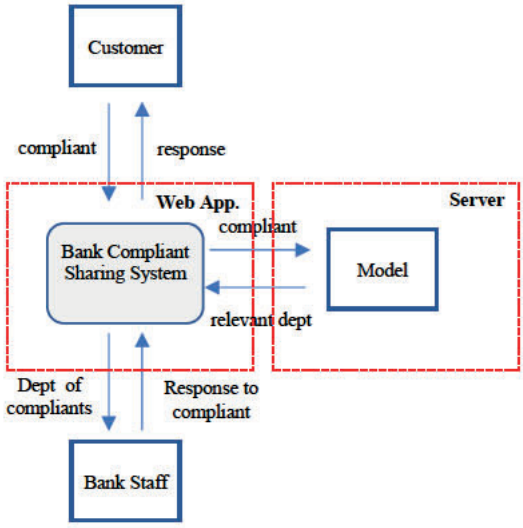


Figure 1: Conceptual data flow diagram of proposed method.

Nowadays, many offered bank services are branched into personal and commercial. Financial institutions have great importance, but they may sometimes have conflicts with their customers. Besides the complaints are quite various, customers look for a solution by using mobile application, web page or calling customer service offered by the bank to express their complaints. However, apart from these methods, they can share their complaints on different platforms that can be seen by both bank employees and people who do not have a relation with the bank by using various websites.

GRU and LSTM deep learning models have given successful results in text classification in recent years. Although there is no significant difference in results of GRU and LSTM, GRU is preferred because of its higher accuracy rate. For training, sample complaints in text form and a data set labeled with the department to which the complaint belongs are required. Since,

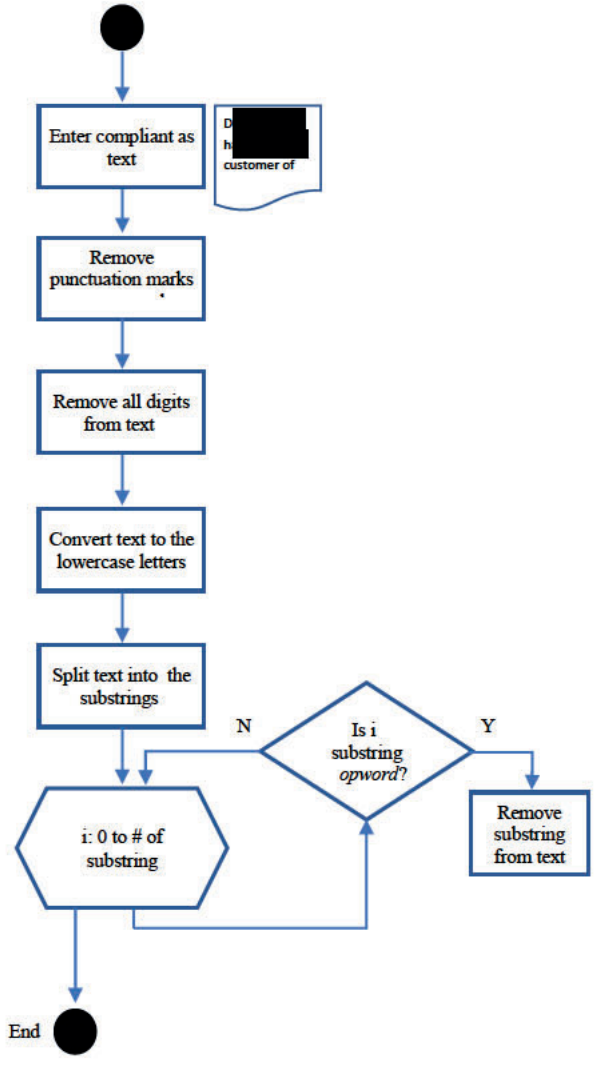


Figure 2: Pre-processing task flow diagram

there is no Turkish data set on this subject, the entire data set was created by us. However, as direct use of customer complaints in raw form may reduce the success rate, it should be passed through a specific pre-processing phase. In pre-processing, removing words that do not effect to the classification is one of the important process that most affect the performance of the model. In order to be accomplish,

Turkish *stopwords* have been found and updated in accordance with the financial sector. Then, extracted words have been given to the model using a proper word embedding method. In the training phase of the deep learning model, each complaint in the dataset is pre-processed and arranged that does not affect the classification. Pre-processing task of proposed method is shown in Figure 2. As we can observe from figure, the punctuation marks, special characters and numbers are removed from the complaint text, respectively. All characters in the complaint are converted to lowercase letters. The complaint text is divided into words and it is checked whether each word is a *stopword*. If the word is *stopword*, it is removed from the complaint. At the end of the transaction, the issued complaint is returned.

The complaint from the user is received in text format and goes through the pre-processing stage. With the tokenization process, they are divided into complaint words and appropriate numbers are assigned to these words in order to give them as input to the neural network. The complaint is given to the model and the department of the complaint is obtained in the output of the model. At the end of the process, the department information is returned. Flow diagram of proposed method is given in Figure 3.

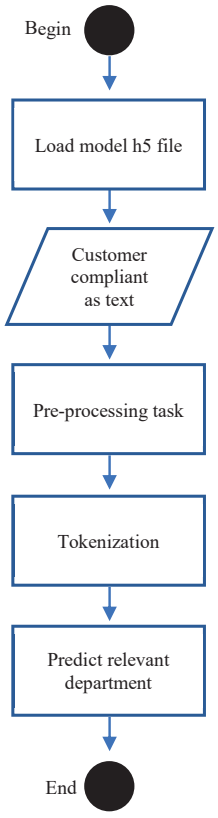
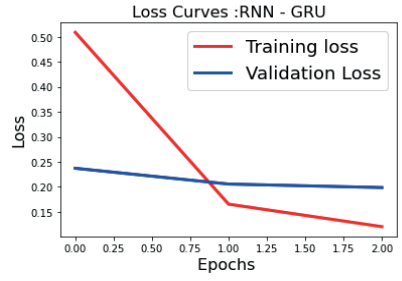


Figure 3: Flow diagram of proposed method

III. EXPERIMENTAL RESULTS

In order to perform the training and testing of the deep learning model all dataset have been generated by us. Raspberry Pi Zero W provides model and web interface communication over an API with RESTful Web Service running on Raspbian, a Linux-based operating system. Raspberry Pi have been used as Python server. Since the application that will serve on the client will communicate through the API, it has been implemented using web technologies such as HTML, CSS and JavaScript. A complaint received through the web application is sent to the server in JSON format. The server analyzes the complaint in JSON format, determines which department or in other words which class the complaint belongs to, and sends it back to the web application in JSON format. Thus, the complaint received from the user is presented back to the user as classified with our model running on the Python server. MySQL has been used as a database management system. In the application, there are two types of users: customer and bank employee. Employees in the web application cannot delete a complaint and report it as resolved. These transactions can only be made by the customer that sharing the complaint. Solved complaints are not deleted from the database. Bank employees and customers with profiles in the system can comment on a complaint. More than one comment can be made for a complaint. After a customer deletes a complaint, the complaint and its comments, if any, cannot be accessed, and are deleted from the complaint database. If the customer updates the complaint as resolved, anyone who visits the website can view the complaint and the comments made to the complaint. The proposed model is able to divide the complaints in text format at the rate of 95% accuracy. There are 2 applications running on the server of the study. Python is used to use the deep learning model and Java is used to access the database. On the client side, TypeScript programming language is used.

The deep learning model has been implemented using the Python programming language. Then, it was run on a server and used in the web application. The dataset is a 40377-line file with *.csv* extension that holds complaints and department informations. There are 5 classes in the data set. Figure 4 shows the Loss-Epoch graph and accuracy value of the model developed with GRU.



Accuracy: 95.27%

Figure 4: Loss-epoch graph of GRU model

The deep learning model data consists of 40377 lines (90% train and 10% test). Learning rate is select 0.01, number of epochs is 3 and batch size is 128. Softmax is selected for activation function and 20% dropout used. Categorical crossentropy loss function and Adam optimizer are used.

For each class, TP (True Positive), FP (False Positive), accuracy values is calculated and ROC curve is drawn. Figure 5 shows the performance values and ROC curves of the classes.

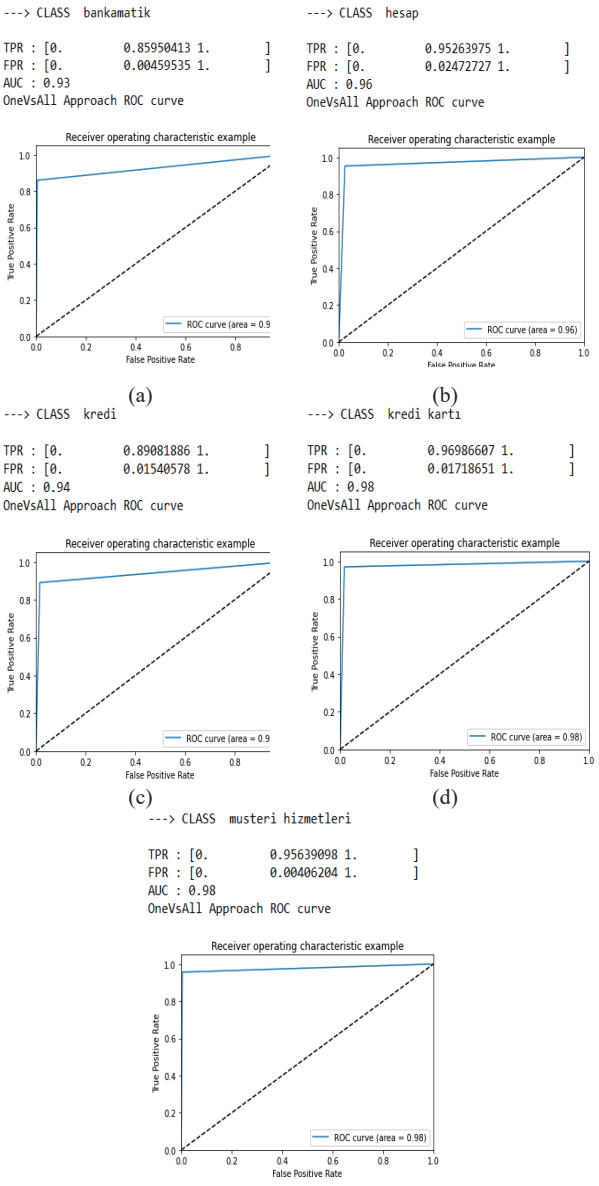


Figure 5: Performance values and ROC curves for cash and account classes a) bankamatik b) hesap c) kredi d) kredi kartı e) musteri hizmetleri

IV. CONCLUSION

The proposed deep learning model can divide complaints into different departments with an accuracy of 95% rate for 5 classes: ATM, account, credit, credit card and customer service. The model can detect the department that a complaint belongs to in less than 1 second. While sharing the complaints in the web application, if the user does not trust the result detected by the model, the customer can choose the department to which the complaint will be submitted. Thus, complaints are forwarded to the relevant department of the bank in a short time without the need for human intervention. Bank employees can also respond to complaints by using the application or increase customer satisfaction by communicating with the customer. The deep learning model developed can classify 5 different departments. When it is desired to make complaint classification for more departments, the model can be retrained by updating the data set.

REFERENCES

[1] M N Alabay, "Müşteri Şikayet Yönetimi", Uluslararası Yönetim İktisat ve İşletme Dergisi, cilt 8, sayı 16, sayfa 137-157, Haziran 2012
[2] https://www ing com tr/tr/ing/hakkimizda/musteri-sikayet-yonetimi
[3] Ç Acı, A Çırak, "Türkçe Haber Metinlerinin Konvolüsyonel Sinir Ağları ve Word2Vec Kullanılarak Sınıflandırılması", Bilişim Teknolojileri Dergisi 12 (3), 219-228
[4] Y Santur, "Sentiment Analysis Based on Gated Recurrent Unit", 2019 International Artificial Intelligence and Data Processing Symposium (IDAP), Malatya, Turkey, 21-22 Sept 2019
[5] S Caner, M Erten, "Dağıtık Hizmet Reddi Saldırılarının Tekrarlayan Sinir Ağlarıyla Tespiti", 2019 Ulusal Yazılım Mühendisliği Sempozyumu, Eylül 2019
[6] J Gong, Q Teng, H Zhang, L Du, S Chen, M Z A Bhuiyan, J Li, M Liu and H Ma, "Hierarchical Graph Transformer-Based Deep Learning Model for Large-Scale Multi-Label Text Classification", IEEE Access, Vol 8, pp 30885 - 30896, 2020
[7] K S Tai, R Socher and C D Manning, "Improved semantic representations from tree-structured long short-term memory networks", Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing, Vol 1, Beijing, China, 2015
[8] Y Kim, "Convolutional neural networks for sentence classification", 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP 2014), Doha, Qatar, 2014
[9] J Liu, W C Chang, Y Wu and Y Yang, "Deep learning for extreme multi-label text classification", Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2017), pp 115–124, August 2017
[10] X Zhang, J Zhao and Y LeCun, "Character-level convolutional networks for text classification", in Advances in neural information processing systems, vol 1, pp 649–657, 2015
[11] Y Kim, Y Jernite, D Sontag and A M Rush, "Character-aware neural language models", Thirtieth AAAI Conference on Artificial Intelligence, pp 2741–2749, 2016
[12] J D Prusa and T M Khoshgoftaar, "Designing a better data representation for deep neural networks and text classification", 2016 IEEE 17th International Conference on Information Reuse and Integration (IRI), Pittsburgh, PA, USA, 28-30 July 2016
[13] A Conneau, H Schwenk, L Barrault and Y Lecun, "Very deep convolutional networks for text classification", Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics, Valencia, Spain, Vol 1, pp 1107–1116, April 2017
[14] K Simonyan and A Zisserman, "Very deep convolutional networks for large-scale image recognition", in 3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings, 2015
[15] K He, X Zhang, S Ren and J Sun, "Deep residual learning for image recognition", in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2016
[16] H T Le, C Cerisara and A Denis, "Do convolutional networks need to be deep for text classification?", in Workshops at the Thirty-Second AAAI Conference on Artificial Intelligence, 2018

[17] G Huang, Z Liu, L Van Der Maaten and K Q Weinberger, "Densely connected convolutional networks", in Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, 2017
[18] B Guo, C Zhang, J Liu and X Ma, "Improving text classification with weighted word embeddings via a multi-channel TextCNN model", Neurocomputing, Vol 363, Pages 366 - 374, 21 October 2019
[19] L Mou, R Men, G Li, Y Xu, L Zhang, R Yan and Z Jin, "Natural language inference by tree-based convolution and heuristic matching", Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics, Vol 2, pp 130 - 136, Berlin, Germany, August 2016
[20] L Pang, Y Lan, J Guo, J Xu, S Wan and X Cheng, "Text matching as image recognition", in 30th AAAI Conference on Artificial Intelligence (AAAI 2016), pp 2793-2799, 2016
[21] J Wang, Z Wang, D Zhang and J Yan, "Combining knowledge with deep convolutional neural networks for short text classification", IJCAI International Joint Conference on Artificial Intelligence, pp 2915 - 2921, August 2017
[22] S Karimi, X Dai, H Hassanzadeh and A Nguyen, "Automatic Diagnosis Coding of Radiology Reports: A Comparison of Deep Learning and Conventional Classification Methods", pp 328–332, Vancouver, Canada, 2017
[23] S Peng, R You, H Wang, C Zhai, H Mamitsuka and S Zhu, "DeepMeSH: Deep semantic representation for improving large-scale MeSH indexing", Bioinformatics, vol 32, 12, 2016
[24] A Rios and R Kavuluru, "Convolutional neural networks for biomedical text classification: Application in indexing biomedical articles", in BCB 2015 - 6th ACM Conference on Bioinformatics, Computational Biology, and Health Informatics, 258-267, September 2015
[25] M Hughes, I Li, S Kotoulas, and T Suzumura, "Medical Text Classification Using Convolutional Neural Networks," Studies in Health Technology and Informatics, 235:246-250, 2017

Information technology in healthcare: Architecture of healthcare information systems

S. CAGLAR[1], A.A. ALTUN[2], F. KARA[3], and A. TAGHIYEV[2]

[1]Ministry of Health, President of the Community Health Center, Aksaray-Sultanhani/Turkey,
[2]Selcuk University, Konya/Turkey, anart@selcuk.edu.tr
[3]Ministry of Health, General Directorate of Public Health, Ankara/Turkey,

Abstract **– Many healthcare institutions are only now exploring the possibilities of Apache Hadoop&Spark distributed processing architectures. The manuscript describes the options for the architecture of building an information system based on cloud technologies: centralized and distributed systems for healthcare. The positive and negative aspects of these options are presented, based on which the conclusion is made about the higher prospects of the implementation of distributed architecture. The basic requirements for the development of identification control systems, access, and standards for the construction of electronic health records are formulated.**

Keywords **– Big data technology, cloud computing, healthcare.**

I. INTRODUCTION

At present, information, and communication technologies, as a highly demanded field of knowledge, are developing in a variety of directions. Among the paradigms of development of modern information technologies, we can highlight the direction called Cloud Computing (*CC*). This direction is a rapidly developing promising stage of modern computer science. *CC* ideology is to transfer the organization of calculations and Big Medical Data (*BMD*) processing [1-4] to a significant extent from personal computers to the servers of the World Wide Web. The concept of cloud structures is based on the belief that the Internet is able to meet the needs of users in generating and processing of *BMD* in a wide range of their needs [5-7]. Among the principles of creating the Healthcare Information System (*HIS*), listed in the concept of creating *HIS*, which significantly affect the architecture of the system can be identified the following: (a) Ensuring interoperability of various Information Systems (*IS*). (b) Creation of Applied Information Systems (*AIS*) on model Software as a Service (*SaaS*). (c) Decision-making a to modernize the inherited and develop new components of the *HIS* [7], taking into account the maximum possible preservation of existing software and hardware based on the analysis of the total cost. When creating a *HIS*, it is most expedient to use *CC,* which is rapidly developing now and is widely used to implement complex geographically distributed *IS* [5]. In healthcare, *CC* is defined as follows: *CC* is a service delivery model that enables healthcare institutions to access a pool of shared resources across the board, efficiently and on demand. The main advantage of using cloud solutions is a significant increase in the efficiency of automated medical processes [6] and a reduction in the cost of creating, maintaining and developing of *IS* and *IT* infrastructure due to the following artifacts [8]: (a) centralization of *IT* resources; (b) virtualization of *IT* resources; (c) dynamic management of *IT* resources; (d) ubiquitous access to *IT* resources; (e) automation of *IT* processes; (f) simplification of *IT* services; (g) standardization of *IT* infrastructure.

Based on the approaches offered by cloud technologies, we can identify two possible options for building the architecture of *HIS*, which call centralized and distributed architecture.

II. CENTRALIZED ARCHITECTURE

In the centralized architecture (see Fig.1) almost all components *AIS* are placed in the centralized core of the system. The core of the system is built based on several technological platforms united by the corresponding telecommunication infrastructure. The technological platform is either an own separate Data Processing Center (*DPC*) or a place leased in the *DPC* by one of the operators-a specific numbers of racks on which the server and communication equipment is located. In order to ensure the reliability of the *HIS*, it is necessary to have at least two geographically separated technological sites.

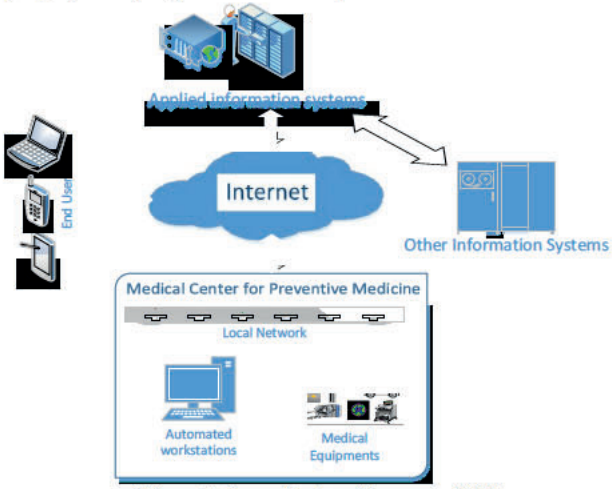


Figure 1: Centralized architecture of *HIS*.

The centralized architecture assumes that access to and work with applied software from Medical and Preventive Healthcare Institutions (*M&PHI*) is based on the SaaS model. In this case,

only a web browser is required to work with *AIS*. Local Area Network (*LAN*) is being built in *PHI*, to which automated workstations (*AW*) of various specialists, as well as medical equipment (ME) are being connected. *LAN* is connected to the Internet via appropriate communication equipment to work with the core of the *HIS*. *BMD* from *ME* is uploaded to the centralized core storage, also via the Internet. It should be noted that this data storage architecture, *BMD* is not processed in *M&PHI*, except for computerized *ME*, where the manufacturer of *ME* provides such data storage. It is possible to provide access and work with *HIS* applications for authorized end users, including mobile users connected via the Internet, which is especially relevant for the emergency medical care system and medical teams providing emergency assistance. Personal computers, laptops, tablets, and smartphones can be used to work with *HIS*. Access to *HIS* for a wide range of users can be implemented through the so-called Personal Office (*PO*) with the possibility of using electronic keys, with the appropriate procedure of authorization and authentication by the requirements of information security. It is possible to organize the interaction with other external *IS*, including those created as part of the implementation of the program of electronic public healthcare.

A. *Benefits of the centralized architecture*

- Reduce the cost of acquiring and maintaining *LAN* servers in medical and healthcare institutions and their maintenance by technicians.
- Possibility to organize single monitoring of the entire system.
- Storage of *BMD*, in particular archives and medical images, obtained in various from *M&PHI* and available to physicians from other institutions (availability of primary data, not only medical reports).
- Availability of all necessary medical information about the patient for remotely consulting physicians and doctors-experts of the Compulsory Medical Insurance (*CMI*) system.
- Protection of Data Storage Systems (*DSS*) in a single place.

B. *Main challenges of the centralized architecture*

- High requirements for reliable connection of *M&PHI* to the Internet and reliable functioning of the access network itself. Disruptions in the functioning of the access network in the centralized architecture lead to the complete stoppage of the work of *M&PHI*. Therefore, the minimum requirement is that *M&PHI* should be connected to the Internet via two physical channels with different routes.
- Sufficiently high bandwidth and bandwidth stability requirements between *M&PHI* and the core of *HIS*. For easy work of the employees of *M&PHI* with the number of workplaces 10-20, the connection speed should be at the level of 10 Mbit/s and the bandwidth of the through the channel from the workstations to the centralized application not less than 256 Kbit/s. However, the problem is that today, it is impossible to guarantee the required level of service provision for the end-to-end channel because traffic from the *AW* to the application usually passes through the networks of several operators.
- The use of Decision Support Systems (*DSS*) [7] and information and reference systems in a multi-user online mode will further increase the requirements for communication channels.
- Organization of crypto protection of transmitted data.
- Difficulties (often-actual impossibility) in the use of already implemented local *IS*.
- Using the SaaS model for specific medical applied software for all types of medical institutions will require the deployment of multiple specialized *IS* in the core.
- Major problems with connecting private institutions to the centralized system, which are unlikely to give up their own local *IS*.
- Actually dependence on one manufacturer (a group of manufacturers for different components) at the realization of the uniform *AIS* in the centralized kernel.

III. ARCHITECTURE OF DISTRIBUTED SYSTEMS

Distributed *HIS* is a system with web-based applications that are accessible via the Internet and that share both local and global resources (see Fig.2). In the distributed architecture, as well as in the centralized architecture, components of *AIS* are placed in the centralized core of the system. At the same time, however, local *IS* are preserved (still functioning) in the *M&PHI*, possibly without implementation of some functions; for instance, without archival storage of electronic health records (*EHR*) of retired patients, without storage of medical images (especially large volumes, such as *MRI*, etc.), backup of current information (registered or hospitalized patients), etc. Local *IS* in this architecture serve as a kind of intermediate storage (as if a cache) between the centralized core and local *AW*. The local *IS* database is a kind of replication of a part of the centralized database. Local *IS* should be entrusted with auxiliary functions that do not require centralization, for instance, the provision of work with systems to support medical decision-making [4].

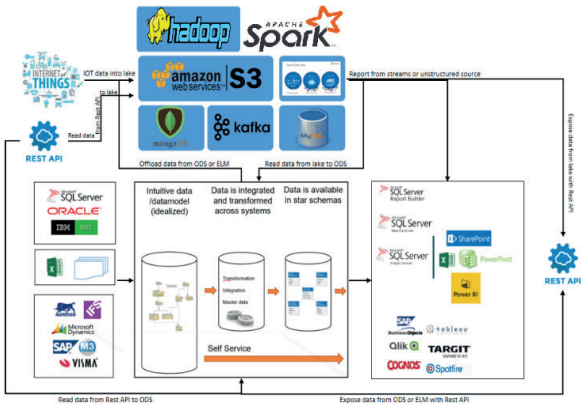


Figure 2: Distributed architecture of *HIS*.

Data exchange between the local *IS*, and the centralized core can be carried out on request or schedule. The challenges of modern health care require important decisions. Distributed systems can help to ensure that *BMD* is processed in real time and that analysis results are available and effective so that researchers can study the genetic architecture of cancer cells;

health care professionals can monitor intensive care patients [1, 2], i.e., they can obtain valuable information, manage costs, and provide patients with better treatment options through the analysis of *BMD* in a distributed system environment [4, 5]. The work of users, including mobile users, connected via the Internet, external *IS*, and users through *PO* is possible only with a centralized core. It means that when a person contacts the *M&PHI* that he or she has not previously visited, the medical information available to him or her will have to be uploaded into the local *IS* from a centralized repository. If in the course of providing medical care the information has been changed or supplemented, these changes should be uploaded back to the centralized database.

A. The benefits of distributed architecture

- Substantial reduction of requirements for reliable connection of *M&PHI* to the access network, bandwidth and stability of the through channel bandwidth between *M&PHI* and the core of the system. The *M&PHI* remains operational even in case of complete disconnection from the access network.
- Ability to use previously implemented local *IS*, including those integrated with radiological and laboratory systems, including auto-analyzers, with which exchange from the cloud is virtually impossible.
- Possibility to connect to a centralized system of private *M&PHI*, including the definition of the volume of *BMD* stored in a centralized system.
- Maintaining competition between manufacturers of *IS* and thereby improving the quality of systems.

B. The main challenges of distributed architecture

The most critical tasks in the implementation of a *HIS* with a distributed architecture:

- Development of a personal identity management system.
- Development of the concept and standards.
- Development of an access control system.

Patient Identity (*PI*) management system of *EHR*, implemented at the federal level, should provide: (a) Issuance of new identifiers (based on personal universal electronic identification cards); (b) maintenance of current identifiers (storage of *PI* and attached data, including exact date and place of birth, etc.), which will provide a single point of access to information about the patient's state of health; (c) Archive of identifiers (deceased persons or persons who left the country); (d) Issuance of temporary identifiers for contract visitors from other countries; (e) Receipt of all variable personal information to be stored in conjunction with the *PI* assigned at birth.

The development of a concept and standards for *EHR* should provide: (a) Implementation of a single modular architecture of *EHR*, including mandatory, auxiliary modules for all types of medical institutions. Mandatory modules should contain vital data, diagnoses, information on the treatment of chronic diseases, services provided, life, and disease histories. Auxiliary modules may include information that has a limited time for general use, such as emergency care and intensive care. (b) Person-centric data integration at the federal level or in the regional segment of *HIS*, which will result in physical or virtually merged records of the health status of patients, regardless of the place of first data recording and storage. (c) Use of all-Turkish classifiers for both questionnaire and medical modules. (d) Organization of information exchange based on open or agreed protocols and formats. (e) Conversion of *EHR* from databases of various institutional *IS* at the regional or federal level. (f) Formation of integrated *EHR* based on consolidation into approved fixed modules (for permanent and temporary integration).

The development of a system for managing access to *EHR* involves: (a) Maintaining links at the federal level to *EHR* stored in regional data centers. Maintenance of references to EMI at the federal level, stored in regional *DPC*. (b) Authentication and authorization of a medical worker before accessing a patient's *EHR*. (c) Authorized access to sections or modules of *EHR* depending on the position held and functional duties performed at that time (for instance, duty doctor). Modular exchange of information will speed up its transmission/receipt and reduce the load on communication channels. (d) Request to *DPC* of regional segments of *HIS* to view/receive/send medical data by reference in the federal segment. (e) Mandatory registration of the patient's rights is a permit to transfer *BMD* to other doctors (except in cases of unconscious state or court decision on the patient's incapacity).

IV. Conclusion

The current manuscript presents a prospective study of a distributed system in healthcare. Due to the lack of computing power for processing *BMD*, there is a need for modern distributed *HIS*, which reduce the costs, time and inefficient use of resources, that is, to solve challenges it is necessary to develop a distributed *HIS* based on *CC* [5]. We understand that even despite various objective difficulties, *CC* will be actively developed, including, of course, in the market of medical and healthcare software. As *CC* develops, the situation is likely to continue to evolve.

Along with the implementation of new projects in healthcare, it is necessary to study and manage the risks of *IT* projects [3, 8], that is, in healthcare, there is a strict rule, which has been verified for many years of medical practice is not harmful. When implementing distributed *HIS* using *CC*, existing *IS* should be taken into account. At the same time, the transition to the new principles of building an information meta-system should be based on the introduction of a single *EHR*, formed based on a multi-module structure. It should take into account the needs of a wide range of users at all levels of the *HIS*, including *CMI*. In the case of *CC*, this means mandatory accounting, risk assessment, and management of *IT*, and the use of time-tested specialized solutions (for instance, Microsoft solutions) [8, 9].

References

[1] A TAGHIYEV, A A ALTUN, et al , "A Machine Learning Framework to Identify the Causes of HbA1c in Patients With Type 2 Diabetes Mellitus," *Journal of control engineering and applied informatics,* vol 21, no 2, pp 34-42, 2019, DOI: 10 6084/m9 figshare 11980896 v1

[2] A TAGHIYEV, A A ALTUN, and S CAGLAR, "A Hybrid Approach Based on Machine Learning to Identify the Causes of Obesity," *Journal of control engineering and applied informatics,* vol 22, no 2, pp 56-66, 2020

[3] A ALTUN, S CAGLAR and A TAGHIYEV, “Modern approaches to the application of machine learning and data mining methods in the healthcare industry,” I*nternational Conference on Engineering Technologies (ICENTE'19)*, Konya/Turkey, pp 35-38, 2019

[4] A A ALTUN, S CAGLAR and A TAGHIYEV, “Big medical data processing with apache spark in healthcare,” *11th International Congress of Social Sciences with Current Research,* pp 1-4, 2019

[5] A ALTUN and A TAGHIYEV, “Big Data Analytics on Cloud Based,” *International Symposium on Human and Social Sciences (ISHSS'18),* https://drive google com/file/d/1hNbtFHv8tNBENEmU_8f3ETeES5IQkN-1/view, Antalya/Turkey, pp 19, 2018

[6] A A Altun and A Taghiyev, “Advanced image processing techniques and applications for biological objects,” 2017 2nd IEEE International Conference on Computational Intelligence and Applications (ICCIA), Beijing, 2017, pp 340-344, doi: 10 1109/CIAPP 2017 8167235

[7] A ALTUN and A TAGHIYEV, “Big Data Design: Semantic Structures and Causal Models for Decision Making,” *International Symposium on Human and Social Sciences (ISHSS'18),* Antalya/Turkey, pp 19, 2018, https://drive google com/file/d/1hNbtFHv8tNBENEmU_8f3ETeES5IQkN-1/view

[8] A ALTUN and A TAGHIYEV, “Risk Management in IT Project,” I*nternational Conference on Engineering Technologies (ICENTE'17)*, Konya/Turkey, pp 253-258, 2017

[9] A TAGHIYEV, “Microsoft Big Data Solutions,” Big Data Baku 2015, Baku/Azerbaijan, http://cedawi org/docs/Anar-Tagiyev-Big-Data-Day-Baku-2015--Microsoft-Big-Data-solutions pdf

Analysis of Flight Delay Propagation with Bayesian Networks

Bulent SOYKAN[1] and Serpil EROL[2]

[1]Old Dominion University, Norfolk/VA/USA, bsoyk001@odu.edu
[2]Gazi University, Ankara/Turkey, serpiler@gazi.edu.tr

***Abstract* - Flight delays are becoming more prevalent in airline operations, costing airlines billions of dollars each year. Most flight delays in airline operations are unpredictable and stochastic in nature and are the result of things such as: late arriving aircraft, air carrier delay, national aviation system delay, and extreme weather. These flight delays can be mitigated by embedded schedule buffer times. To model the complex and stochastic airline schedule system, different optimization and simulation models are employed. In this paper we investigate the use of Bayesian Networks to analyze flight delay propagation.**

***Keywords* - Airline Delay Analysis, Bayesian Networks, Flight Delay Propagation.**

I. Introduction

Flight delays' negative impacts are comprehensive, increasing flight delays pose a major threat on the whole air travel system and cost airlines at billions of dollars each year. As stated in research made by Airlines for America (A4A), in 2012 it is estimated that delays caused $7.2 billion in direct operating costs for scheduled U.S. passenger airlines, and considering that crew costs are estimated at $16.26 per minute, the total crew related cost is approximately $1.5 billion [1]. By reducing delays, airlines can reduce operating costs and increase revenue. Modeling techniques can be used to analyze current systems and to investigate new operating approaches.

Flight delay propagation occurs when late arrivals at an airport cause late departures, which in turn cause late arrivals at the destination airports. Delay propagation is difficult to analyze using traditional methodologies, because it is inherently a stochastic and dynamic phenomenon. Therefore, it must be characterized by probability density functions, and statistical or dynamical analysis methods. In order to model flight delay propagation, simulation models should be employed. The advantage of this approach is its capability of dealing with complex systems as well as catching the stochastic and dynamic nature of the system. Also, the simulation model allows us to observe how the whole system behaves given the interactions between fixed flight schedules and stochastic disrupting events emerging in operations [2]. In this paper, we investigate the use of Bayesian Networks as a method to analyze flight delay propagation.

The paper is organized as follows: Section II discusses the Literature Survey for this problem. Section III explains the Methodology used for analyzing flight delay propagation. The results and observations are discussed in Section IV. Discussion, Conclusion and Future works are stated in Section V and VI, respectively.

II. Literature Survey

In order to model and simulate the flight delays and flight delay propagation, a variety of methodologies has been adopted. Abdelghany et al. proposed a deterministic model to predict the propagated delays along aircraft route based on network optimization models [3]. Beatty et al. presented a 'delay multiplier' metric to estimate the scale of delay propagation [4].

Schaefer and Millner proposed a detailed policy assessment tool to model delay propagation in a network of airports when facing inclement weather conditions. To limit flight delays, proactive schedule recovery models are developed [5].

Lui and Ma studied the impact of delay propagation on flight delays in busy airports by using Bayesian Network. Departure delay is chosen as the object variable in the proposed model, Netica software package is used to establish Bayesian Network model that is based on available data and experts' experience. And it is concluded that the propagation is more notable and severe when the arrival delay is moderate or high, and propagation leads to flights' departure delay. The primary aim of the model is to reduce the influence caused by delay-propagation from busy airports, which are international hub airports that can relieve the flight delays of whole aviation system. However, the model takes into account only one busy hub airport as the research object, whose delay problem is very serious, which is actually not realistic. Also, the influence of flights' delay-propagation in different states is described intuitively not analytically [6].

Even though different approaches have been employed to model the flight delays and flight delay propagation, the stochastic effects of delays resulting from various delay causes have not been thoroughly captured.

III. Methodology

The methodology of this research was focused on achieving some analysis of flight delay propagation using Bayesian Networks. Through the use of real-world flight delay data and GeNIe software, we were able to draw some conclusions.

A. Flight Delay Propagation

Flow of flight delay demonstrated in Fig. 1 illustrates the relationship between departures, arrivals, and delays of two consecutive flights. The solid lines with arrows show the planned flight and the dotted ones display the actual flight, for flight legs *i* and *j*. *PDT*, *ADT*, *PAT* and *AAT* are the Planned Departure Time, Actual Departure Time, Planned Arrival Time and Actual Arrival Time, respectively. The Planned Turn Time between flights *i* and *j* (PTT_{ij}) is the time between *PAT* of flight *i* and *PDT* of flight *j*, see Equation (1). *ATT* (Actual Turn Time) is the ADT of flight leg *j* minus the *AAT* of flight leg *i*. If *ATT* is smaller than Minimum Turn Time (*MMT*) the flight is considered delayed. It is assumed that the same resources (airplane and crew) are assigned to flight *i* and *j*. As shown in Fig. 1, when flight leg *i* is delayed, the Slack time (red parathesis) between two flights completely and some part of the *MMT* is consumed, and this causes Propagated Delay (*PD*). Hence, the Total Departure Delay (*TDD*) of flight *j* is consist of the PD_{ij} and the Independent Departure Delay (*IDD*) of flight *j* itself. Likewise, the Total Arrival Delay (*TAD*) of flight *j* comprises PD_{ij} and the Independent Arrival Delay (*IAD*). Although *IDD* includes only the independent delay before a flight departs, *IAD* includes both *IDD* and the additional independent delay in the air or at the origin and destination airports.

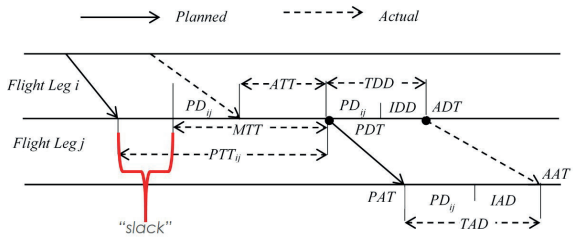


Fig 1 Flow of propagated delay

$$PTT_{ij} = PDT_j - PAT_i \quad (1)$$

$$Slack_{ij} = PTT_{ij} - MTT_{ij} \quad (2)$$

As shown in Fig. 1, when flight leg *i* is delayed, the Slack time ($Slack_{ij}$) between two flights is consumed completely and also some part of the *MMT* is consumed, and this causes Propagated Delay (*PD*). Hence, the Total Departure Delay (*TDD*) of flight *j* is consist of the PD_{ij} and the Independent Departure Delay (*IDD*) of flight *j* itself. Likewise, the Total Arrival Delay (TAD) of flight *j* comprises PD_{ij} and the Independent Arrival Delay (*IAD*). Although *IDD* includes only the independent delay before a flight departs, *IAD* includes both *IDD* and the additional independent delay in the air or at the origin and destination airports. We can calculate the following relationships in order to find the propagated delay:

$$TDD_j = max(ADT_j - PDT_i, 0) = max(PD_{ij} + IDD_j, 0) \quad (3)$$

$$TAD_j = max(AAT_j - PAT_i, 0) = max(PD_{ij} + IAD_j, 0) \quad (4)$$

$$PD_{ij} = max(TAD_i - Slack_{ij}, 0) \quad (5)$$

Consequently, for each day of operation, we can obtain *PD*s using *IAD*s, which assumes that the first flight of each day has zero propagated delay.

B. Data

The analysis was generated with a field-collected actual flight data from a small-scale Turkish airline which operates short-haul domestic flights on a hub-and-spoke network flights which consist of 17,200 flights spanning a six-month period. Fig. 1 is an example of the raw data structure. The raw data was not useable in the original format and provided some challenges to organize it sufficiently for the calculation of event probabilities. For example, if a flight departed in the evening and arrived the following day, we had to correct for a negative value derived from a time difference. Additionally, the date formats were not consistent and required code to make them identical, and thus useable.

DATE OF OPERATION	FLIGHT LEG NU.	ORIGIN	DEST.	PLANNED ARRIVAL TIME	ACTUAL ARRIVAL TIME	PLANNED DEPART TIME	ACTUAL DEPART TIME
20130101	THY7001	ADB	ESB	4:45	04:54	06:00	06:00
20130101	THY7002	ESB	ADB	7:30	07:20	8:50	08:56

Fig 1 Example of raw data format

Based on the categories we devised for the model parameters, we were required to compute the probabilities for 184 different queries corresponding to event combinations (e.g., long delay given an afternoon flight on a weekday). We could not account for canceled or diverted flights, because these events were not captured in the dataset.

C. Delays Explained

Delays are arrival-based, meaning the baseline measure for all delays start with the arrival process. Fig. 1 graphically portrays the flow of propagated flight delay.

D. Implementation with GeNIe

The diagram in Fig. 3 depicts our Bayesian network model in GeNIe modeling software. We investigated the impacts of Time of Day, Day of Week, and Season on arrival and departure delays. An advantage of this model is that you can input any similar dataset, for example from a different airline, to analyze the effects of different model parameters.

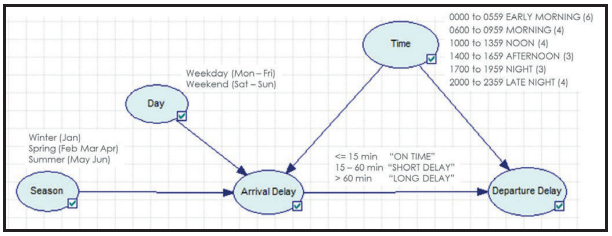


Fig 3 GeNIe Bayesian network model

IV. RESULTS

A. Statistics on Flights per Category

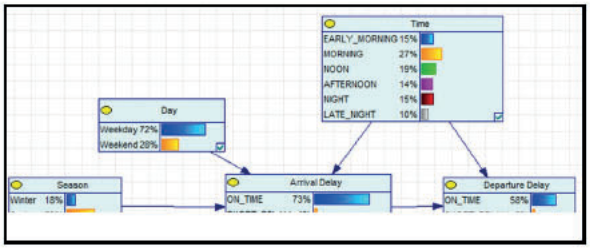


Fig 4 Percentage of flights per category

Fig. 4. shows the percentage of flights for a given category. All things being equal (i.e., no evidence set for a parameter) arrivals were on time 73 percent of the time and departures were on time 58 percent of the time. The season percentages seemed skewed on the outset; however, recall the data was for six months, so accordingly the seasons reflect this. Winter represents on month of data; spring, three months; and summer, two months.

B. Use of Evidence

Evidence is set in the model for a particular parameter. There are many queries that could be performed by setting the evidence for the model. We investigated a fraction of the 184 combinations available in this particular model. Fig. 5 shows the results of running the model with evidence set to a long delay for an arrival delay. This situation results in no on-time departures and a high percentage of long departure delays.

The example in Fig. 6 shows the evidence for time of day set to afternoon. The results tell us that flights operating in the afternoon only departed on-time about half the time.

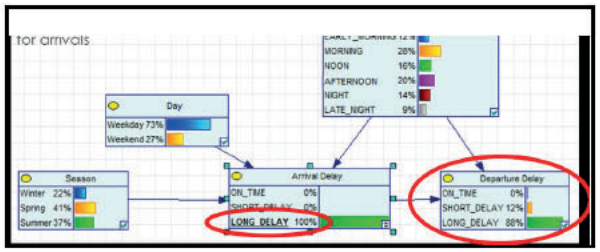


Fig 5 Evidence set to long delay for arrival delays

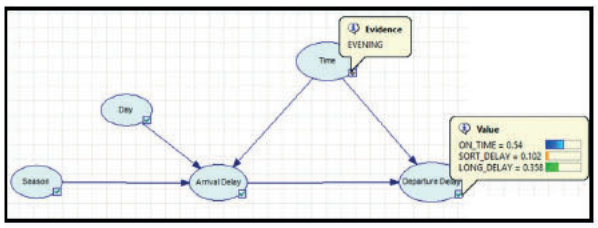


Fig 6 Evidence set to afternoon for time of day

C. Delays Based on Time of Day

The results can include as many statistics as you can query, so we selected some results to demonstrate the kind of output achieved in this model. The chart in Fig. 7 shows the probabilities of arrival and departure delays based on the factors we considered. The following graphs in Fig. 8 and Fig. 9 provide a visual inspection of the data that puts into perspective the importance of time of day as relates to propagated delays.

		ARRIVAL DELAY	DEPARTURE DELAY
EARLY MORNING	ON TIME	.75	.62
	SHORT DELAY	.07	.06
	LONG DELAY	.18	.33
MORNING	ON TIME	.72	.56
	SHORT DELAY	.04	.08
	LONG DELAY	.24	.36
NOON	ON TIME	.77	.65
	SHORT DELAY	.04	.09
	LONG DELAY	.19	.26
EVENING	ON TIME	.63	.54
	SHORT DELAY	.06	.10
	LONG DELAY	.31	.36
NIGHT	ON TIME	.75	.56
	SHORT DELAY	.05	.11
	LONG DELAY	.21	.33
LATE NIGHT	ON TIME	.78	.57
	SHORT DELAY	.01	.13
	LONG DELAY	.21	.30

Fig 7 Delays based on time of day

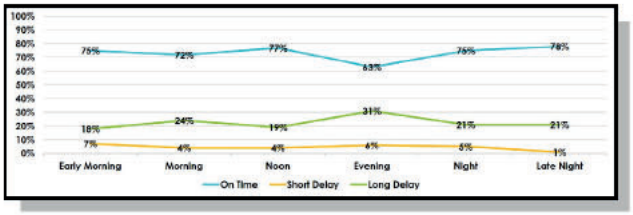


Fig 8 Graph of arrival delays based on time of day

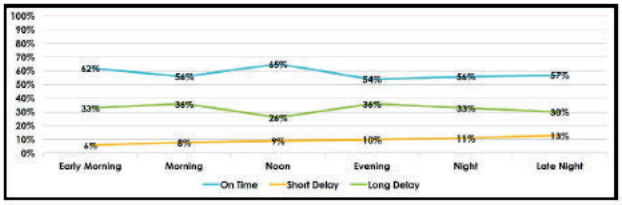


Fig 9 Graph of departure delays based on time of day

D. Key Observations

Through viewing different parameter combinations through the queries, we identified some key observations.

- If a flight ARRIVES on time, then most likely it will DEPART on time.
- If a flight ARRIVES with a short delay, there is an equal chance of a short DEPARTURE delay
- If a flight ARRIVES with a long delay, there is almost no chance it will DEPART on time.
- As flights progressed throughout the day, short DEPARTURE delays increased.
- Generally, flights operating in the middle of the day experienced the highest on-time ARRIVALS and DEPARTURES.

V. DISCUSSION

Intuitively we expected that delays would build up or propagate throughout the day. What we discovered is that this is generally true for short delay departures; however, short delays did not propagate into long delays. We suspect this is due to the built-in slack time (or buffer time) into the schedules and the pilots' ability to increase speed in some circumstances especially with the presence of a tailwind.

VI. CONCLUSION

Flight delays are unavoidable in real-world airline operations, because we can almost never predict with certainty what will happen in the future that may cause a delay. But we can minimize the effects of delays by investigating the causes of delay propagation and by eliminating these causes. In this paper, we utilized Bayesian Networks to determine the propagated delays based on prior knowledge and specific circumstances.

We expected that flights would propagate delays as the day progressed, which is true for short delay departures; however, short delays did not propagate into long delays. We suspect this is due to the built-in slack time and the fact that pilots can make up time in the air with increased air speed and presence of a tail wind.

Future work includes modeling weather related delays and combining Bayesian Networks with traditional statistical methods, which can be used to provide more accurate modelling and therefore more precise prediction. Also integrating expert judgment with statistical analysis in structure construction and parameter estimation can not only save time and effort, but improve prediction accuracy as well.

REFERENCES

[1] A f America (2013) *Annual and Per-Minute Cost of Delays to U.S. Airlines* Available: http://www airlines org/Pages/Annual-and-Per-Minute-Cost-of-Delays-to-U S -Airlines aspx

[2] C -L Wu, "Inherent delays and operational reliability of airline schedules," *Journal of Air Transport Management,* vol 11, pp 273-282, 2005

[3] R Beatty, R Hsu, L Berry, and J Rome, "Preliminary evaluation of flight delay propagation through an airline schedule," *Air Traffic Quarterly,* vol 7, pp 259-270, 1999

[4] K F Abdelghany, S S Shah, S Raina, and A F Abdelghany, "A model for projecting flight delays during irregular operation conditions," *Journal of Air Transport Management,* vol 10, pp 385-394, 2004

[5] L Schaefer and D Millner, "Flight delay propagation analysis with the Detailed Policy Assessment Tool," in *Systems, Man, and Cybernetics, 2001 IEEE International Conference on*, 2001, pp 1299-1303 vol 2

[6] Y -J Liu and S Ma, "Flight delay and delay propagation analysis based on bayesian network," in *Knowledge Acquisition and Modeling, 2008. KAM'08. International Symposium on*, 2008, pp 318-322

Social Campus Application with Machine Learning for Mobile Devices

Mücahit TÜNEL[1] and Şaban GÜLCÜ[2]

[1]Necmettin Erbakan University, Konya/Turkey, mucahit.tunel42@gmail.com, Orcid id: 0000-0003-3471-9235
[2]Necmettin Erbakan University, Konya/Turkey, sgulcu@erbakan.edu.tr, Orcid id: 0000-0001-7714-8861

Abstract **- In this study, Social Campus Application which can be used on mobile device is explained. Different types of communication have been developed throughout human history. With the rapid change of technology, today's popular communication organs have been the internet and social media. Social media can cause us to spend most of our time on the phone. Social media platforms that offer many options such as quick access to news, instant communication and sharing are indispensables among many of us. This study is about a social media application that is intended to be used among university students. This application was designed which users can be informed of the activity created, open topics and make comments about them, and shop through the market. The application is android-based and contains fp-growth algorithm in the market section. The fp-growth algorithm is an association rule algorithm used to find frequent item sets. The fp-growth algorithm offers the user who supplied together with the analysis of what products the most. The application includes the messaging property between users and this property is based on the technology of firebase, node.js and socket.io. Django and sqlite was used in the web section and the application was developed using the react-native technology.**

Keywords **- Android, React-Native, Fp-Growth algorithm, Social Campus Application, Social Media.**

I. Introduction

THE ever-increasing ideas of societies such as storing, sharing and using information have led people to new searches. Undoubtedly, one of the most events of 20th century is the discovery of the internet and it introduction to the public service. With the advent of the Internet, people have provided easy, cheap, fast and secure access to information in many areas [1, 2]. In this way, people started to produce new productions. This new communication phenomenon and process, which emerged from the combination of communication satellites and developing technology after the discovery of computer technology, caused our age to be called the "Information Age" [3]. Thanks to the Internet, we can easily send and receive information from one end of the world to the other, and with the advancing technologies in many areas, we can easily do things that we used to have difficult in the past, on the computer. In the information age, information is power and it is necessary to use this power effectively. Even carefully sharing the existing information with someone else is one of the qualities that can increase this power.

The fact that the internet is an interactive medium and that it enables everyone to express their opinions without limitation is beneficial for communication experts, while unlike other mediums, the internet's duality and suitability for multiple communication draw attention[4]. Providing such a wide area of use and fast communication, the Internet is widely used in universities. With the networks created within the university; Connection via wireless networks is provided to students, academic and administrative staff and employees upon request [5].

Social media is a media system that enables one-way information sharing to be achieved through two-sided and simultaneous information sharing. Social media is a field of application that allows the sharing of information, different views/opinions and experiences by public websites and quickly embeds the internet world in our lives [6]. People can perform many operations such as mutual conversation, messaging and information sharing in two directions, especially thanks to social media applications. Instead of the term social media, which has turned into a very effective structure, terms such as social network, social web, social networking sites are sometimes used both in daily life and in academic literature [7]. Social media meets needs such as communication, entertainment, spending time or following the agenda. But besides these, it can provide an environment for information pollution. Social media, which respond to the social demands of people from all cultures and all walks of life, causes an increase in information sharing and adds a new dimension to socialization [8]. Thanks to social media, polls can be created, people's opinions can be taken, their comments on a subject can be learned, and as a result, new developments and new studies can be realized.

Today, many social media applications have emerged and these applications have played very important roles in communication. Many people around the world use social media applications actively. According to the data, active social media usage in the world is 3.81 billion. According to Statista data, this situation is presented graphically in Figure 1. [9].

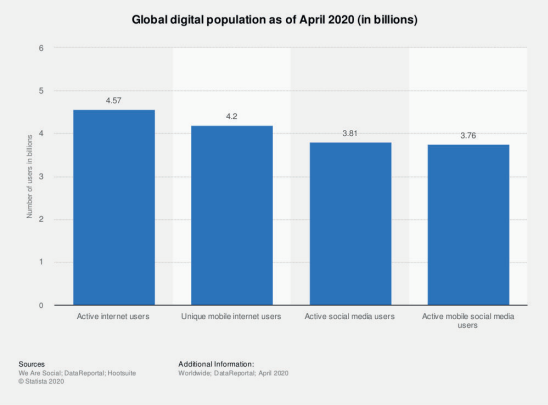


Figure 1: Global digital population as of April 2020 [9].

As you can see, most of the people who use the internet use social media actively. According to Statista data, the most used social media platform is Facebook. This situation is presented graphically in Figure 2 [10].

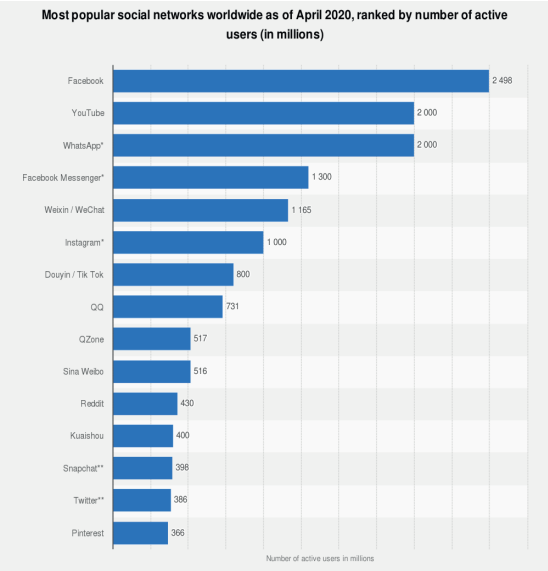


Figure 2: Most used social networks in the world (million)[10].

Social media usage is also very common in our country. It is seen that many people from almost all walks of life use social media. When we look at the rate of social media usage in our country, there are differences in social media usage according to the usage in the world. The Youtube application in our country ranks first with 90% according to the data of the third quarter of 2019. According to Statista data, this situation is presented in Figure 3 [11].

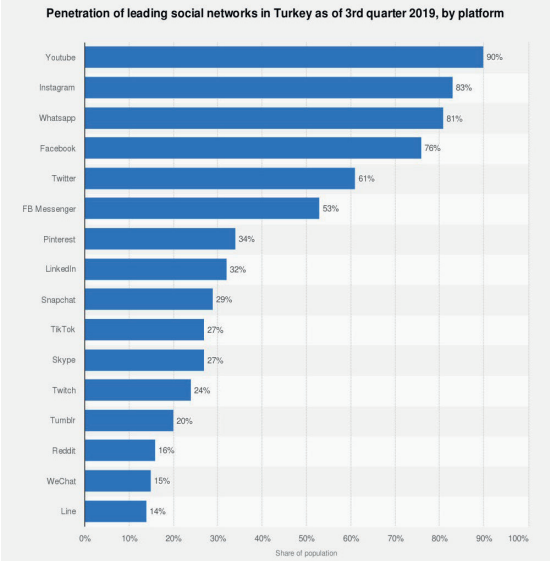


Figure 3: 3rd Quarter Year 2019 in Turkey Social Media Usage Rates [11].

In this study, due to the use of social media so much, an application that includes many features was developed. The social campus application is a social media application that mostly targets university students. It is a social media application where new events can be created and people can be informed, opinions can be received on any subject when there are events such as concerts, trips, and people can reach each other and do business thanks to the market page. However, the people who use the application have the ability to message among themselves. The main purpose of this application is to inform users about the events, to learn their comments and opinions, and to facilitate trading.

Social media applications with many different features are developed for individuals to use according to their needs. There are many social media applications for messaging, sharing or informational purposes. Some of these are given below.

The Uno Social application is an application that allows you to connect with your classmates. It provides access to features such as events, calendars or maps. It offers the services offered on campus. The application has more than 1000 downloads so far.

The Mobile Campus application allows users to reach the online campus on the go. It offers the opportunity to follow active lessons or old lessons. Social groups can be established and notes can be shared. There is also a web version of the application. Features that cannot be done with the mobile application are available on the web page. The application currently has more than 5000 installs.

Facebook is a social media application that aims to enable people to communicate with each other and share information. It has a very wide user area. It is suitable for 13 years and older. Local shopping is provided thanks to the Marketplace. Games can be played inside. Photo, video can be shared and instant notifications can be received. It currently has more

than 5,000,000,000 downloads.

WhatsApp is a free instant messaging application. It has the ability to send messages, calls, photos, videos and audio messages. Since the session is always open, instant notifications can be received. It also has the ability to share the instant location. It currently has more than 5,000,000,000 downloads.

The social campus application developed in this study is an application that encourages users to socialize and provide information exchange. It informs us in advance of the events to be organized. By opening topics, opinions of other users are received. It covers everyone, not a specific university community, and is user-friendly. The products that are wanted to be sold with the Market section are delivered to other users by posting ads.

Our study is divided as follows. Detailed information about the subject is given in Chapter 1 and previous studies are listed. The technologies used and the developed system are explained in Chapter 2. Results and outputs are presented in Chapter 3, and the conclusion part is shared in Chapter 4.

II. Materials and Methods

A. Association Rule Mining and Fp-Growth Algorithm

Association rule algorithms [12] is a method used to find relationships between variables in large databases. Finds the possibilities of realization according to the data in the database. Predictions about the future are made within these possibilities.

Fp-Growth algorithm is one of the fastest approaches used to find frequency data on the data set [13]. It works fast especially on large datasets. Fp-Growth works with divide and conquer logic [14]. Performs 2 scans on the database. In the first scan, the support values of all objects are calculated. In the second scan, a tree data structure is created. Thus, success can be achieved in large data sets in only two scans. The pseudo code of the algorithm is presented in Figure 4.

Procedure *FPgrowth**(T)
Input: A conditional FP-tree T
Output: The complete set of all FI's corresponding to T.
Method:
1. **if** T only contains a single branch B
2. **for each** subset Y of the set of items in B
3. output itemset $Y \cup T.base$ with count = smallest count of nodes in Y;
4. **else for each** i in $T.header$ **do begin**
5. output $Y = T.base \cup \{i\}$ with $i.count$;
6. **if** $T.FP\text{-}array$ is defined
7. construct a new header table for Y's FP-tree from $T.FP\text{-}array$
8. **else** construct a new header table from T;
9. construct Y's conditional FP-tree T_Y and possibly its FP-array A_Y;
10. **if** $T_Y \neq \emptyset$
11. call *FPgrowth**(T_Y);
12. **end**

Figure 4: Fp-Growth Algorithm Pseudo code [15].

B. Used Technologies

React-Native:

React-Native [16] is an open source mobile application development library produced by Facebook in 2015. Since it is used with React components, improvements can be made to the web, android and ios, the interface of which is completely created by ourselves. Finally, version 0.63.0 is available. Development is done using Javascript. When the codes are compiled, they turn into a native structure. In React-native, great attention should be paid to the life cycle of the application in order to increase its performance. Because doing important operations such as database connection in the wrong places does not give the desired image by creating instant changes on the application screen.

Django:

Django [17] is a high-level Python web application framework. It was first released on July 15, 2005. Finally, version 3.0.7 is available. It has a simple use. It has a very fast, secure and highly scalable structure. Django uses the Model-View-Controller (MVC) structure. It offers an easy development environment thanks to the many methods it offers out of the box. It uses object-relational matching between relational databases. Thanks to serialization processes, data in XML and JSON format can be used. By default, the SQLite database is used.

Firebase:

Firebase [18] was founded in 2011 by Firebase Inc. produced by. It is a platform designed for mobile and web application development. Finally, the 7.15.0 version is available for Javascript. It contains many features such as real-time database, cloud messaging, cloud storage. It has a useful architecture to process application data in real-time. Firebase was acquired by Google in 2014 and has been integrated with many Google services.

Socket io:

Socket.io [19] is a javascript library prepared for use in real-time applications. It works with Node.js. Socket.io provides real-time analytics and bidirectional communication. Finally, version 2.3.0 is available. As of version 2.0, Socket.io uses µWebSockets as the underlying WebSocket library.

III. Developed System

The social campus application was made to enable users to be aware of the events created, to start a topic and to learn different comments about it, and to facilitate shopping transactions through the market. Add, delete and edit options are available for each section. Thanks to the search page, the desired topic, event or user can be found. Popular topics and market products are added to favorites. Users can message among themselves. It has the feature of blocking each other. When blocking is done, a structure has been created in which old messages can be seen but new messages will not be seen.

The social campus application was developed using the

0.61.5 version of React-Native. Django version 2.2.1 and SQLite database are used. Database operations were performed using the object model with Django. Chat operations were performed using Firebase 6.7.1 version. The connection between React-Native and Django is made using rest technology. Data transfer is made with JSON. Instant communication is provided by using Socket.io version 2.3.0. In the market part of our application, the Fp-Growth algorithm, which is an association extraction algorithm, is used. In this way, you can see which products are mostly examined together with the product you have examined.

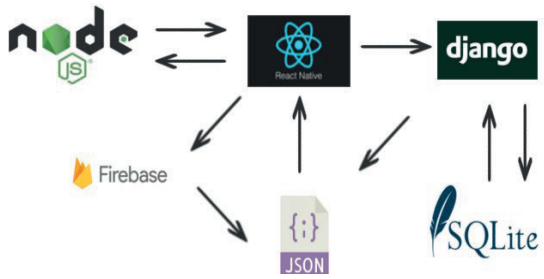


Figure 5: System Architecture.

In order to use the application, you must first register using your gmail address. The user can access the contents after logging in with the registered account. You can list the events, follow the opened topics or create a new topic, enter the market section and contact users for their necessary needs. Interface shapes of the application are indicated in Figure 6, Figure 7 and Figure 8.

The screens shown in Figure 6 are the first screens the user will encounter in the application. First, you have to register and then log in. After login, you will first see the event list screen. Details can be reached by clicking.

Other pages in the main plan are shown in Figure 7. Products are listed on the Market screen. The options at the top are filtering operations. The topics that have been opened can be listed on the second screen. By clicking it, the comments within it can be listed. The last screen is the search screen. There are 3 different search options.

Figure 8 shows the profile page first. Settings are made here. The second and third screens are the section where the event and product details are listed, respectively. Product deletion and editing operations can be performed on the market screen.

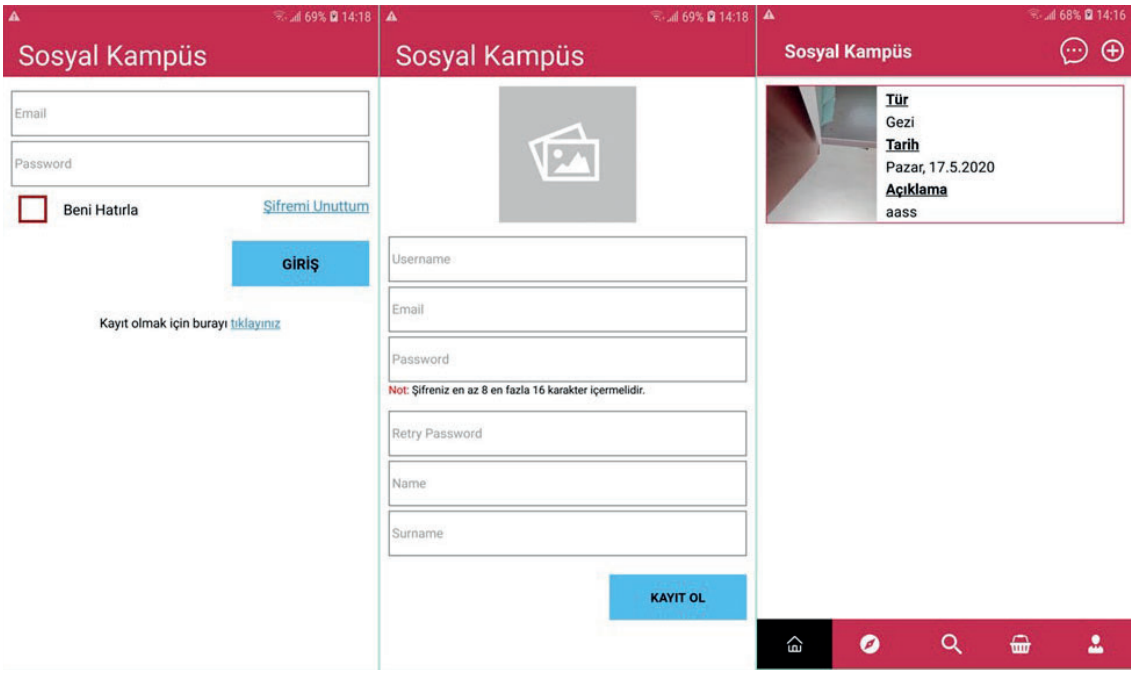


Figure 6: Login-Registration-Homepage Screens.

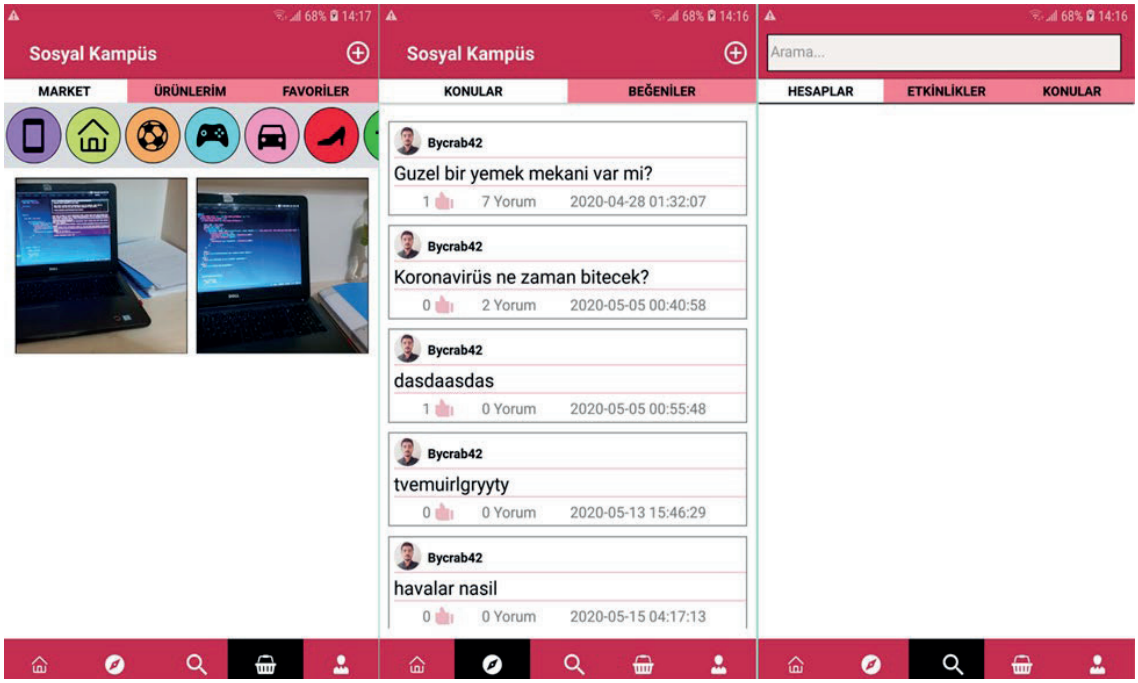


Figure 7: Market-Subject-Search Screens.

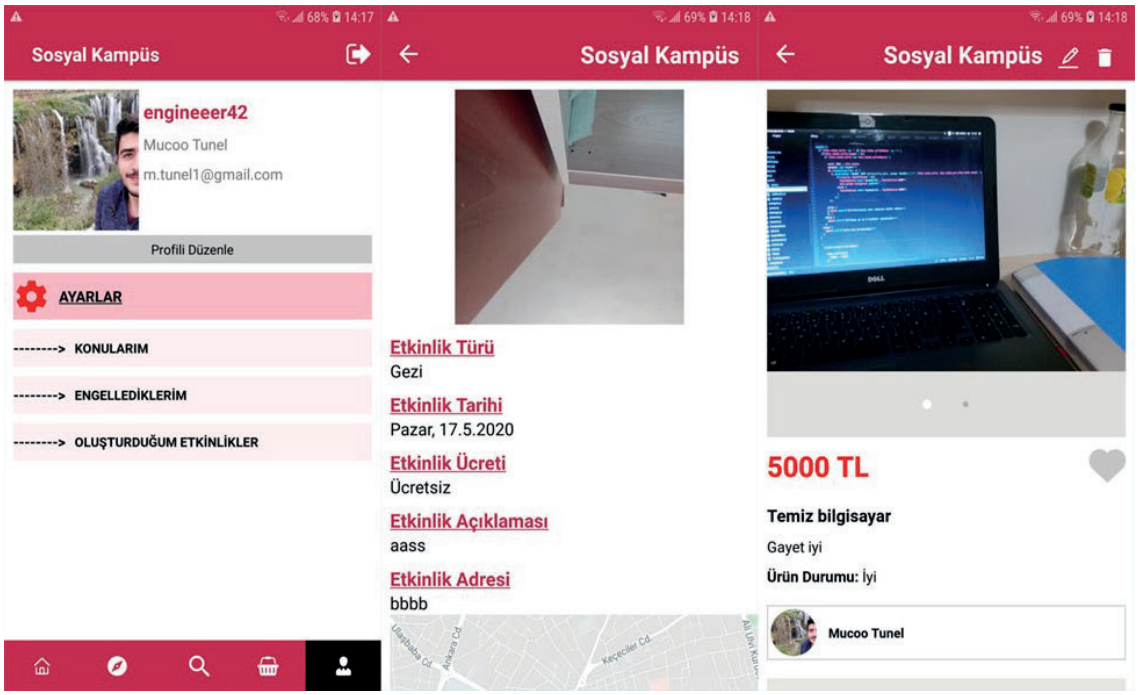


Figure 8: Profile-Event Detail-Market Detail Screens.

IV. Conclusion

In this study, a social media application was developed to inform the users and to get their comments and opinions, generally covering university students. Creating and publishing an event, creating a topic, listing the topics and receiving comments and posting ads were successfully completed. The Fp-Growth algorithm has been successfully shown to the users, along with the products that users have viewed, along with the most listed products.

The application system is working successfully. In particular, the codes of the application have been optimized more legibly, resulting in high performance in terms of speed. User-friendly design has been created. Thus, users can easily use the application and find the menu they are looking for easily.

Currently developed application works on android devices. It does not work on ios operating system. The next work is to develop the application to run on the ios operating system.

REFERENCES

[1] H Tutar, M Yılmaz, Genel iletişim kavramlar ve modeller (6 Basım) Ankara: Seçkin Yayıncılık, (2008)

[2] Ş Gülcü, M Tünel, Mobil Cihaz için Apriori Algoritması Tabanlı Kitap Özetleri Uygulaması, Ejons International Journal on Mathematics, Engineering - Natural Sciences, 3 (12), (2019) 46-58

[3] A Aziz, İletişime Giriş, Aksu Kitapevi, İstanbul, 2008

[4] J Chipchase, A Theaker, İnternetin Halkla İlişkiler Alanında Etkili Bir Şekilde Kullanımı, Murat Yaz (çev), İstanbul: Kapital Medya Hizmetleri AŞ, (2006) 357-391

[5] M Yılmaz, Üniversitelerde halkla ilişkiler: Kurumsal web sayfaları ve sosyal medya uygulamaları üzerine bir değerlendirme, Selçuk Üniversitesi Sosyal Bilimler Enstitüsü, 2015

[6] T Weinberg, The new community rules: Marketing on the social web, (2009)

[7] İ Sayımer, Sanal ortamda halkla ilişkiler, Beta, 2008

[8] A Okay, A Okay, Halkla İlişkiler Kavram Stratejileri ve Uygulamaları, İstanbul: DR Yayınları, 2012

[9] Digital population worldwide, https://www statista com/statistics/617136/digital-population-worldwide/, 2020

[10] Global Social Networks Ranked By Number of Users: https://www statista com/statistics/272014/global-social-networks-ranked-by-number-of-users/, 2020

[11] Turkey Social Network Penetration: https://www statista com/statistics/284503/turkey-social-network-penetration/, 2020

[12] R Agrawal, R Srikant, Fast algorithms for mining association rules: Proc 20th int conf very large data bases, VLDB, 1994, pp 487-499

[13] J Han, J Pei, Y Yin, R Mao, Mining frequent patterns without candidate generation: A frequent-pattern tree approach, Data mining and knowledge discovery, 8 (2004) 53-87

[14] H Li, Y Wang, D Zhang, M Zhang, E Y Chang, Pfp: parallel fp-growth for query recommendation: Proceedings of the 2008 ACM conference on Recommender systems, 2008, pp 107-114

[15] G Grahne, J Zhu, Fast algorithms for frequent itemset mining using fp-trees, IEEE transactions on knowledge and data engineering, 17(10), 1347-1362, 2005

[16] React-Native: https://reactnative dev/, 2020

[17] Django: https://www djangoproject com/, 2020

[18] FireBase: https://firebase google com/, 2020

[19] Socket io: https://socket io/, 2020

New Approaches to Improving Software Code Quality

O.MUTLU[1] and O.KALIPSIZ[1]

[1] Yıldız Technical University, İstanbul/Turkey, ozge.mtlu@hotmail.com
[1] Yıldız Technical University , İstanbul/Turkey, okalipsiz@gmail.com

Abstract - **The criteria used to evaluate the software have a big impact on the goal, so decisions to be made about which metrics will be more efficient based on the needs of the software are very important. Our goal is to measure the complexity of the software at the design stage for two reasons: Complexity is an important factor affecting the quality of the software in many aspects such as reusability, understandability and maintenance cost. Measuring complexity in the design phase can provide many advantages in quality, due to the contribution of this phase in reducing the cost and effort of redesign and sustainability.**

In this study, software quality and metrics were examined, and different calculations of code complexity, which is one of these metrics, were discussed, applied and answers were sought for the question of how to make it more efficient. Applications are made in database objects that do not have large quality tools like object oriented programming languages. As a result, it has been proven that the complexity is measurable and controllable in these objects, and it has been seen that a more detailed and extensible model can be used successfully than the existing complexity calculations.

Keywords **- Code quality, software quality, code complexity, maintainability**

I. INTRODUCTION

While trying to respond to the rapidly developing and changing needs in the software industry, it is of great importance to have reliable, cost-effective, standards-compliant and error-free software. Because the effect of errors can be wide, it can cause big problems, which can cause negativities such as endangering human life, stopping or preventing communication and transmission, faulty production, customer loss or bankruptcy. It is a testing method to avoid all these problems. Testing is known as a method of obtaining information about the functionality and quality characteristics of the software and identifying errors. However, most errors encountered / discovered during the testing phase are late findings. Because every error in the test in the production environment or even before causes the SDLC process to be rerun from the previous stages and must be fixed and retested respectively. This means a serious waste of time and money. The early detection of software-related errors and the detection and elimination of these errors during the development phase provide significant savings. The aim is to improve software quality by preventing and correcting errors during software development. [1]

Quality software should be easy to maintain, easy to understand, well structured and open to development. In addition, the error rate per line of code and the error rate during the operation period should be at the lowest level. According to the ISO 9126 quality standard, it is said that the product of the software is updateable in summary for the ease of maintenance feature. These updates; error correction in software, software improvement and adaptation of software according to changing and developing needs. These changes, which are related to the fundamentals of the software, bring the ease of maintenance feature to a very important point. If a software application cannot well support changes due to ease of maintenance, maintenance and repair work becomes quite expensive. [13]

Complexity makes it difficult to read and understand the software and therefore change it. It is among the main reasons for the errors. Considering all this, we can say that the complexity is the opposite of software quality. Therefore, it plays an important role to measure complexity and keep it within certain limits. In this study, after showing the importance of complexity metric, its role in improving software quality, and the advantages of measuring complexity in the design phase, we will focus on applications that measure software complexity.

The basic information used in Chapter 2 is explained. Information was given about software quality, its relationships with software metrics were examined, and complexity, one of the metrics that most affected the code quality, was emphasized. It has been explained how the proposed complexity models are calculated, and their advantages and disadvantages are compared. Section 3 is the section where the relevant study is detailed. The complexity models proposed so far have been tested on database objects and the results have been interpreted, and a more comprehensive and expandable model that includes all these models has been tried to be created. A general judgment is made by comparing the results of the study evaluated in Chapter 4 with the literature.

II. LITERATURE REVIEW

Software quality is a planned series of activities to ensure that quality is integrated into software. If we define it, quality is the degree to which the software or process meets the specified customer or user needs, expectations and requirements. . Quality activities should be followed at every stage of the software life cycle in order to be efficient. For each step, it

must be ensured that the resulting output and ultimately the product meet the quality standards at the end of the process.
A key element in controlling, managing and improving the software development process is software measurement. Throughout the analysis, design and coding stages, the outputs of product development must be measured, monitored and managed so that they can be verified against predetermined criteria. Product malfunctions can be prevented by making software measurements. These metrics warn developers and engineers of development errors, thereby preventing errors and defects, both before and during the product release, and also help monitor software development.

Software projects can be digitized using software metrics to make measurements. Thus, by comparing the data obtained in existing models, information about the status and progress of the software can be obtained and new models can be developed.

Software Quality Metrics
In order to objectively interpret and determine software quality, its features need to be digitized. Various sets of metrics are used for this. These sets of metrics are divided into class and method level. [2]

Class-Level Software Quality Metrics
It is a set of metrics introduced by Chidamber-Kemerer in 1994, abbreviated as CC. The metrics calculated according to this approach are as follows;

1. WMC (Weighted Method per Class): WCM is the number of weighted methods in a class.
2. DIT (Depth of Inheritance Tree): It is the depth of the inheritance tree, ie DIT is the distance of the furthest path from a class to the root element in the inheritance tree.
3. RFC (Response for a Class): It is calculated by adding the number of methods triggered by the class in a class and the number of methods inherited.
4. NOC (Number of Children): The number of subclasses. It can be defined as the number of children of the class or the number of classes derived from a class.
5. CBO (Coupling between Object Classes): It is the dependency between object classes. While there is no inheritance between classes, the use of properties or methods in one class in another class indicates dependency between classes. That is, if class A uses the attributes or operations of class B, then class A is said to be coupled to class B. CBO is the number of different classes to which a class depends other than inheritance, it is used to measure productivity and reusability.
6. LCOM (Lack of Cohesion in Methods): It is the compatibility of methods. For each attribute, the percentage of the methods using that feature is determined and averaged, and this value is subtracted from 100. The value obtained is the LCOM value. While the intra-class cohesion should be high, the coupling between classes should be low. [1-3]

Software Quality Metrics at Method Level

1. Number of Lines of Source Code (Lines Of Code - LOC)
This metric was originally developed to estimate the man-hours of a project. LOC is commonly represented as: Lines of Code (LOC) counts each line, including comments and blank lines.

2. Sustainability Index
The cyclic complexity calculates, based on the combination of halstead volume, the number of source code lines and the number of comment lines, and a score of how maintainable a program is.
In a nutshell, it is the metric that measures the complexity of maintenance of a module. Thus, it shows how easy it will be maintained. It is abbreviated as MI. It uses values such as row number, cyclomatic complexity and halsread completiy when performing the calculation.

The formula used in calculation is as follows;
MI= 171- 5.2 *ln (v) - 0.23 * (G) – 162 * ln (LOC)
V= Halstead Complexity
G= McCabe Cyclomatic Complexity
LOC = Line Of Code

3. Complexity
The metric that we will go into details in our study measures the difficulty of the module in making decisions. The nature of the software always involves some complexity, but this level must be kept under control. As the complexity increases, the understandability and sustainability of the software decreases. The software becomes unusable at the point where complexity overcomes quality. The models suggested so far are as follows;

A. Cyclomatic Complexity
Cyclomatic Complexity is a software metric. In 1976 Thomas J. McCabe, Sr. It is developed by and used to indicate the complexity of a program. It provides quantitative measurements of the complexity of the program through the source code of a program, based on more information / control flow regardless of program size.

The control flow shows a program as a graph consisting of Nodes and Edge. In the graph, Nodes represent processing tasks while edges represent the flow of control between nodes. We can summarize the complexity as the number of independent paths on this flow. An independent path is defined as a road with at least one side that has not been crossed on any other road before.

e: number of edges, arrows connecting nodes
n: number of nodes
p: number of exit points
cyclomatic complexity;
V(G) = E – N + 2P

Advantages:
• Can guide the testing process. Cyclomatic complexity can be effective in determining the level of the Quality Assurance team's testing. If the complexity coefficient is high, we can conclude that the piece of code / functionality is error prone and requires an in-depth test.
• It gives information about whether the unit tests written by the programmers are sufficient or not, because it reveals all the paths of a method. It helps us focus more on the revealed paths. Developers can be sure that all paths have been tested at least once
• It gives information about how effective, simple and sustainable the code written is.
• Since it will show the code snippets that are complex, it is possible to improve these points and increase the understandability of the code.
• Can be used as a measure of quality
Disadvantages:
• It is a measure of the control complexity of programs, not the complexity of the data.
• In this, nested conditional structures are more difficult to understand than non-nested structures. [3,6,12]

B. Halstead Metrics
Introduced by Maurice Howard Halstead in 1977, this method is one of the complexity measurement methods, which is classified as a measure of complexity, that is, that includes both functionality and software. He argues that by summing their numbers arithmetically, he measures the complexity of software and obtains numerical information about codes.
Using Halstead operators and operands, it extracts the following metrics;

n1 = Unique operator number
n2 = unique number of operands
N1 = Total number of operators
N2 = Total number of operands
N = N1 + N (Length)
n = n1 + n2 (vocabulary)
The above length and word capacity play a role as input in the calculation of Halstead's other metrics. Other metrics are as follows;

Program Length (Length / N)
Represents the sum of the total number of operators and operands. It should be at most 300. It is recommended to redesign a code block with a length value greater than 300. A code block with a length of more than 500 must be reviewed and redesigned.
N = N1 + N2

Program Volume (Volume / V)
It is proportional to the total size of the program. It is the memory space required to store the program; so it is represented by bits. It should be at least 20 and at most 1000. A code block with a volume value greater than 1000 needs to be redesigned.
V = N * log2 n

Program Difficulty
It indicates the fault tolerance of the program and should be at most 50. A code block with values greater than 50 needs to be redesigned.
D = (n1 / 2) * (N2 / n2)

Program Effort (Effort / E)
Shows the effort to understand the program. A code block with values greater than 500,000 may be considered poor quality and may need to be redesigned.
E = D * V

Program Time Effort (T)
The time required for programming (T) is shown as a reasonable time to create the Program. Shows the effort to understand the program in seconds. A code block with values greater than 5,000, ie 1 hour and 30 minutes, may be considered poor quality and may need to be redesigned.
T = (E / B)
Here B is the higher number ranging from 5 to 20. However, B is usually taken as 18 because it gave the best results in Halstead's early experiments.

Program Bug Number (B)
It is the number of errors in the application. The number of errors submitted is related to the overall complexity of the software product. It can be calculated as follows:
B = V / 3000

Advantages of Halstead Metrics:
- Simple to calculate.
- Measures the overall quality of the programs.
- Predicts the error rate.
- Predicts maintenance effort.
- Does not require a full analysis of the programming structure.
- Can be used for any programming language

Disadvantages of Halstead Metrics:
- Depends on the whole code. [7-11]

C. Knot Complexity Account
The first method used to measure software complexity was found by Dijkstra in 1968. In this study, the program quality is defined as a decreasing function of the "goto" commands and the "goto" command is specified as a simple measure of complexity.

To show the flow of the program in Fortran programming, lines with arrows are drawn to show where a jump from one line of text to another line of program text has occurred. With these lines drawn, "GOTO" commands are visualized. The overlapping points are called "knot" or knot. A measure of complexity is obtained from the total number of these knot numbers. [4,5]

D. Oviedo Data Flow Complexity Calculation

The Oviedo complexity measurement method, developed by Oviedo as a "program quality model", takes advantage of the data flow characteristics of a program. According to this approach, a program consists of discrete blocks of sequential expressions with a set of properties. And there is data flow between these blocks. It is executed in sequence.

This model takes control flow complexity and data flow complexity together. The ultimate complexity is the sum of these two. Here, control stream complexity CF represents the total number of edges in the control streamline, while data stream complexity DF represents the total data flow within each block.

Oviedo complexity for a G flow chart;

$$OV(G) = CF + DF$$

While Oviedo is an easy to calculate complexity method, it is not preferred in large programs consisting of large blocks because it takes a lot of time to calculate.

When we group the mentioned complexity models as control flow and data flow, we get a result like the following table. [4,5]

Table 1: Complexity Classification

Control Flow	Data Flow
McCabe	Oviedo
Knot	Halstead

III. RELATED WORK

Banking systems are critical systems in many different ways. A simple rounding error can have a devastating effect on the financial company's reputation, so there is a high pressure on developers to work with precision and test their code as much as possible. However, due to reasons such as competition between banks, legal demands, and meeting the needs of the age, business units demand the rapid implementation of the demands. Developers often, due to time constraints, prefer to reuse solutions that are already working and tested with rapid, minor changes, rather than designing solutions correctly, taking into account the quality and sustainability of their code. This causes rapid development and growth of the system, often based on uncertain and spontaneous decisions, which can result in losing control over the code base in the long run.

At this point, the importance of keeping the code quality and control of store procedures (sp), which are used a lot in the banking system due to the advantages we mentioned above and the intensive database usage, emerges.

Complexity has been calculated on the models proposed so far, evaluations have been made on the usability of the results and a more effective model has been tried to be created.

1.The Fırst Study

The complexity calculation was made using the formula of the Cyclomatic Complexity stated in the literature study.

Step1: As the first step, the key words to be used in calculation are determined. Key Words to be used in our calculation are as follows.

```
string [] keyWordSets = {"END LOOP", "END IF", "ELSIF"};
```

Reading is done from the table with Store Procedures (sp), the contents of each sp are first arranged with the regex library. The editing stages include the following steps;

- A regex statement was written to clear the commented lines. In the following figure, the content of sp (linetext) is given to the expression, and the commented lines are replaced with the "" statement.
- Similar to the above step, a regex statement was written to remove the line where sp was created, and sp was stripped of this expression.

Step 2: It is checked whether each word in the line contains the key words we have determined by returning on the basis of the arranged sp information. If it is, the "idxNode" value is increased by 1, the total number of Edges is increased by 3. And the resulting value is obtained by substituting the obtained parameter values in the formula below.

```
int Result = (totalEdge + idxNode - 1 + 2) - (totalNode + 2) + (2 * 1)
```

The result obtained by giving the complexity level from 1 to 4 is interpreted as 1 if it is between 1-10, 2 if it is between 11-20, and 3 if the complexity is between 21-50 and 4 if it is outside these ranges.

2. The Second Study

Halstead method, which calculates complexity using operators and operands, is used.

Step 1: In the first step, the operands and operators to be used in calculation are determined. The determined lists are as follows;

```
string [] keyWordSetsOperators = {"END LOOP", "END IF", "ELSIF", "=", "-", "+", "/", "*", "++", "RETURN", " <","> "," VARCHAR2 "," NUMBER "," IF "," SELECT "," INSERT "," DELETE "," UPDATE "," COMMIT "," DISTINCT "," WHEN "," CASE " , "THEN", "FROM", "WHERE", "INTO"};
string [] keyWordSetsOperands = {"0", "1", "2", "3", "4", "5", "6", "7", "8", "9", "p_%" "v_%"};

  int n1 = 0; // number of uniq operators
  int n2 = 0; // number of uniq operands
  int N1 = 0; // Total number of operators
  int N2 = 0; // Total number of operands
```

Reading is done from the table with Store Procedures (sp), the content of each sp is first edited by following the steps in the first work with the regex library.

Step 2:

It is checked whether each word on the line contains the operands and operators that we have determined by returning

on the basis of the arranged sp information. If it is, the values of n1 and n2 are increased by 1, and their total number is thrown to the values of N1 and N2.

The result is obtained by substituting the values found in the formula given in the literature study.
Since the store procesure we examined does not contain the term "GOTO", Konot does not fit into the Oviedo complexity calculation as it does not consist of blocks.
The two methods we examined were using the regex library, reading and writing from the database caused performance problems. This led us to do the operations we did in the database.
Although it is observed that the use of Cyclomatic Complexity gives much easier and healthier results compared to the Halstead method, the fact that all keywords affect the complexity of equal value is contrary to practical use. Our study provides the opportunity to give weight to the determined keywords, so more detailed results can be obtained. It can call other store procedures inside. We can add sp's, which we know to work for a long time, to these keywords, assign a certain weight, and move the sps that use it to the top in the order of complexity. Likewise, this method can be used for tables with a lot of data. Thus, we give the user the opportunity to interpret and use complexity more flexibly.

3. The Third Study
Step 1: It is the stage where Store procedures are parsed. The store procedure discussed is first divided into lines, then into words, and the texts to be used in the complexity calculation are determined.
Step 2: The extracted data are mapped here (for example; the beginning and end of the Loop and If conditions, their depths, etc.) and the SP objects and their depths, if any, are kept in 2 different tables.

Depth calculation is made with line information. The first IF or LOOP and END statements found form a block. And their depth is 1. The depths of IF or LOOPs in this block increase by one. For a block that is not inside it, the depth continues to be 1 again.
3. Step: In this step, all data are calculated according to the calculation structure designed parametrically and scored. The parametric calculation structure designed has the following features:
- There is a definition table where the features to be used in calculation are determined. Here, a feature exraction study was performed based on the information obtained in the previous step, and the calculation of the determined features according to their coefficients was also used.
- Another definition table to be used in the calculation is the table containing the control types as the content specified for the parser operation. Here, a study has been made to determine the keywords (control types to be used in the calculation).
- Finally created calculation rules table; It is the table containing the calculation rules, including the cross of the feature definition table and control type definition tables. In this table, the coefficients to be used in scoring for each feature that is desired to be used for each control type, and whether that feature is used or not in that control type are recorded.
- The effect of each calculation rule is set to 1.

Calculation:
From the table where the cross is made, the rules to be included in the calculation are taken respectively, the count of the records that provide that rule in a loop is taken and the effect coefficient of that rule (all taken as one) is added to the general total. It gives the complexity of the total sp we get at the end of all the rules.

IV. Conclusion

In software projects, by measuring the quality, problems that may occur at important points such as maintenance-repair, functionality and reliability can be prevented. Software projects can be digitized using software metrics to make measurements. Thus, by comparing the data obtained in existing models, information about the status and progress of the software can be obtained and new models can be developed.
Complexity is a metric that should always be kept under control, as it is the main reason for poor software quality. In the study, the methods that can be used to calculate the complexity of the different layers in the software architecture, especially the database objects, are emphasized and the effects of the results on the error estimation are analyzed.

References

[1] A.Kıral and T.E. Ayyıldız ," Yazılım Kalite Metriklerinin Kıyaslanması: Örnek Bir Olay İncelemesi", 12. Ulusal Yazılım Mühendisliği Sempozyumu (UYMS' 2018) , İstanbul, Turkey,2018
[2] Ç. Çatal And B. DİRİ, "Yazılım Metriklerini Kullanarak Düşük Kaliteli/Yüksek Riskli Modüllerin Otomatik Tespiti," *1. Yazılım Kalitesi ve Yazılım Geliştirme Araçları Sempozyumu* , İstanbul, Turkey, pp.1-5, 2008
[3] Ayman Madi, Oussama Kassem Zein and Seifedine Kadry, "On the Improvement of Cyclomatic Complexity Metric," International Journal of Software Engineering and Its Applications Vol. 7, No. 2, March, 2013
[4] A.Yıldızhan and A.Çetin, "Yazılımın Bilişsel Karmaşıklığını Ölçme Üzerine Bir İnceleme", Sekizinci Mühendislik Ve Teknoloji Sempozyumu,2016
[5] M. R. Woodward, M. R. Hennel, D. Hadley, "A measure of control flow complexity in program text," Transaction on Software Engineering, Sayfa 45-50, 1979, sayfa 45-50.
[6] Khan, Mohd. Kamran; Ahmad, Faiyaz; Khanum, Mohammadi Akheela," Literature Review On Software Complexity, Software Usability And Software Deliverability.", International Journal of Advanced Research in Computer Science . Mar/Apr2018, Vol. 9 Issue 2, p438-441. 4p
[7] Nikhil Govil, "Applying Halstead Software Science on Different Programming Languages for Analyzing Software Complexity", **4th International Conference on Trends in Electronics and Informatics,** Tirunelveli, India, pp.1-25, 2020
[8] Y. Tashtoush, M. Al-Maolegi, B. Arkok," The Correlation among Software Complexity Metrics with Case Study", International Journal of Advanced Computer Research Volume-4 Number-2 Issue-15 June-2014
[9] T. Hariprasad, G. Vidhyagaran ,K. Seenu, C. Thirumalai," Software Complexity Analysis Using Halstead Metrics ", International

Conference on Trends in Electronics and Informatics (ICEI), Tirunelveli, India, pp.1-25, 2017

[10] https://fatihkabakci.com/article-HALSTEAD_KARMASIKLIGI#:~:text=Fonksiyonel%2Dyaz%C4%B1l%C4%B1m%20karma%C5%9F%C4%B1kl%C4%B1k%20%C3%B6l%C3%A7%C3%BCm%20y%C3%B6ntemlerinden,bir%20yaz%C4%B1l%C4%B1m%20metri%C4%9Fi%20olarak%20bilinir. , last accessed 12 November 2020.

[11] https://www.geeksforgeeks.org/software-engineering-halsteads-software-metrics/ , last accessed 12 November 2020.

[12] Tom Mens ," Research trends in structural software complexity", August 2016

[13] Farhan M Al Obisat, Zaid T. Alhalhouli, Tamador I. Alrawashdeh, Tamara E. Alshabatat,"Review Of Literature on Software Quality", World of Computer Science and Information Technology Journal (WCSIT), Vol. 8, No. 5, 32-42, 2018

Guided Feature Selection and Dimensionality Reduction Method for IDS Improvement in DDoS Attacks

S. AL-HELLI[1] and A. AKBAS[2]

[1] University of Turkish Aeronautical Association, Ankara/Turkey, saif.it.84@gmail.com
[2]Cankiri Karatekin University, Cankiri/Turkey, aakbas@karatekin.edu.tr

Abstract **- With the rapid development of computer technologies, network security has become one of the essential issues today. Distributed denial of service (DDoS) attacks are getting more and more important security threats as the improvements on network speeds create new challenges for traditional intrusion detection systems (IDS) to overcome. Moreover, the IDS systems have to deal with a huge set of data from systems and networks to be monitored and managed effectively in real-time, which imposes an important problem due to the large-scale data with many redundant and non-useful features. In this paper, we present an effective procedure composed of feature selection and dimensionality reduction (FSDR). In the selected subset of features, we applied mutual information (MI) for scoring the whole feature space with the use of empirical cumulative density function (ECDF) to select a subset of features that depend on the uncertainty value in MI. After that, we improved the singular value decomposing (SVD) algorithm to reduce the dimensions of the new space of the features selected in MI and we implemented with the back-propagation neural network algorithm to detect various types of DDoS attacks. Primary experiments are implemented in MATLAB environment. The experimental results show us that our method can select the optimal number of features of two datasets. First, the benchmark of NSL-KDD dataset has reduced the dimensions from 41 to 4 and in addition to a modern dataset of DDoS attacks created by Alkasasbeh [1] and its features are reduced from 27 to 6 with highest classification accuracy is still obtained, after carrying out with 5-folds cross-validation. Our suggested method [2] can efficiently minimize dimensions and diminish their computational overhead without jeopardizing on classification accuracy of attacks.**

Keywords **- Distributed Denial of Service, Intrusion Detection System, Mutual Information, Empirical Cumulative Density Function, Singular Value Decomposing, Dimensional Reduction, Feature Selection.**

I. INTRODUCTION

WITH the swift growth of network technologies, security has become one of the primary issues now. One significant security menace is the Distributed Denial of Service (DDoS) attack. Continuous improvements in network speed bring new challenges to traditional detection methods. DDoS attacks make a victim decline supplying normal services on the internet by overwhelming an extraordinary number of malicious traffic. Attackers do not use security vulnerabilities in a networked system but launch attacks against their availability.

In recent years, DDoS attacks and their effects have increased dramatically. Not only did DDoS attacks get bigger, but they also became more frequent and complex. These attacks have several patterns like overwhelming the server using some exploits weak points in service by using software application, the attacks utilize all available resources on the destination device and that consume the entire available bandwidth for the victim device.

The attacks are trying to use computer resources of original users. In general, it consists of stakeholders' efforts to prevent the functioning of Internet sites and services to be executed efficiently or temporarily often. One such solution to detect the DDoS attack is IDS .The IDS ensures availability by checking traffic against malicious action employing a predefined set of signatures representing different attack behavior .One of the major difficulties with these systems is that it cannot detect suspicious behavior application if none of the predefined signatures match the pattern of data being searched [3]. However, the amount of data that must be monitored and processed for the effective management of real-time networks and systems has become a significant issue in large-scale data. In real time the data comprises several features many of which are redundant and irrelevant. Because these features may impact, the performance of detection DDoS maintains expression features to improve IDS efficiency and performance.

In addition, many anomalies only become visible after analysis and examining data in traffic network from multiple locations. For data analysis, we need a preprocessing stage to clean up the data and convert it to an appropriate format so that effective mining is done. Preprocessing is especially helpful for data consisting of complicated structures and multiple dimensions. Consequently, miniaturization is performed as a preprocessing step into which irrelevant and unimportant sizes are identified and eliminated. It is an important stage of mining data to help improve the performance of learning machines. The main advantage of diminishing dimensions is to improve data classification of data groups improve computational efficiency and improve data visualization. It is primarily because the server cannot keep track of the desired source, it is, therefore, difficult to detect an attack without the specific features of the attack being accurately selected. Machine learning algorithms are not a novel approach to DDoS attacks detection and

accuracy varies according to many data due to the massive traffic on the Internet and rapid growth in the types of intrusions and the multiplicity of methods cybersecurity models demand scalability and durable .in addition, over-fitting is difficulty faced in the machine learning is over-fitted including a part of data[4].

Over-fitting may become a negative influence on DDoS attack detection performance. Increasing the dimension of the data set increases the size of the data region leading to occasional data. Consequently, high dimensions of data lead to problems, making classification with datasets with a high number of dimensions difficult [5]. Therefore, one of the key steps in detecting DDoS attacks and analyzing traffic well is feature selection and dimension reduction. The number of features or variables for a smaller group is adequate to obtain important information for network management purposes. Improvements to our study is focused on selecting the best prominent features of DDoS attack detection to increase accuracy detection and reduction of the complexity of the model. It is worth mentioning that a few studies dealt with identifying features and reducing the dimensions for DDoS attacks.

Dimension reduction and feature selection have an active role in pre-processing phase in numerous machine learning usage. As mentioned prior, a data set consisting of multiple features and a few instances may cause overfitting. Typically, the tracking of suspicious and variable traffic and new attacks may not be disclosed unless an intensive study of traffic and could not be disclosed. These features may not reflect their connection to an impact on a particular type of attack. Moreover the aggregated data sets may include interrelated features that may give better results when incorporated.

Hence, the motivation for this task is to develop a method for feature selection which can detect interrelated features across a spacious range of possible features and could be employed for DDoS attacks detection and also detect new attacks that are damaging to Internet networks and users by exploiting limited examples of attack.

In this research, we have developed a new approach to information theory which proposed as combine between Mutual Information Feature Selection and Singular Value Decomposition Dimension Reduction Method named (FSDR) to estimate the relationship between output class and any input feature. The algorithm owns the following characteristics like It efficiently selects features on some of the first step giving weights to all features using the mutual information and the ranking feature Secondly, we will introduce a new calculation for redundancy to degrade the bias of mutual information by using the empirical cumulative function of uncertainty and keeping values of MI in a closed interval. Finally, we use dimension reduction to eliminate noise, and selecting the least number of features maintain accuracy high was achieved which selects features depending on this methodology it is not required to determine the threshold value for the candidate selection process as required in previous methods and studies.

We developed a system based on anomaly improved to detect abnormalities of DoS attack and DDoS attack using neural network classifier application. In addition to modifying the structure and design of the neural network algorithm to suit the nature of the data set used in the search.

This paper is organized as follows: Section I introduces denial of service attacks and the objectives of this study. Section II surveys the literature for the related works in the area of our research. Section III gives general introduction of DDoS attacks. Our proposed methodology is demonstrated in Section IV along with the technical issues. Additionally, it illustrates the integrated phases about the proposed framework and the algorithm for detecting and algorithm combine for feature selection and dimension reduction. Section V demonstrates and discusses the simulation results presented and analyzed. Finally, Section VI concludes this paper.

II. Related work and Literature Survey

In the literature, we have seen that most of the researchers dealt with methods of detecting the attack, did not elaborate on the features and characteristics and ways to reduce the distance and how to choose. Selection of features is crucial for the performance of detection. Using all features from the data set can cause large memory and disk usage and slows down the detection. Therefore, the aim of feature selection is choosing the most representative features which may have the most discrimination power over the data set. S. Sasikala et. Al. [6] introduced a multi-method of filtering consisting of several stages of extracting feature minimizing dimensions by PCA and they used the feature selection, Correlation based Feature Selection (CFS). The latter stage is a method of making an order of features in the new subset selected by Symmetrical Uncertainty (SU) as specified and classification, before being selected for the final stage by the SVM algorithm to select the efficient subset of the dataset for an examination on the enhancement of detection accuracy and best feature subset selection. The method was applicable to 22 diverse medical datasets.

Liao Qin [7] investigated the difference in user behavior based on a weblog, and they introduced a series of 14-dimensional feature space based on user behavior (normal, attack) to describe properties of user behavior. Their study showed the relative variations and/or similarities in DDoS attackers or regular users. Durability applying three classical data mining classification algorithms are adapted to the detection, which are Naive Bayes, RBF Network, and C4.5. Empirical. Results showed that advised features are proper to differentiate legal users or attackers of the application layer. The results show the correctly classified instances are quite high around 99.98%.

Balkanlı et. al. [3] exhibited the best informative features for a strong detection performance on Backscatter DDoS traffic are located to be: frame. Len, IP TTL, and IP prototypes uses C4.5 classifiers trained on initial attack traffic is possible, by the feature selection techniques, namely Chi-Square and Symmetrical Uncertainty, are applied to build different C4.5 Decision Tree classifiers. The results revealed that the auto-generated rules by the C4.5 Decision Tree classifier using only 7 features could reach more than 95% precision in detecting new Backscatter DDoS attack patterns.

Chaitanya et al. [8] proposed detection method utilizes a set of feature selectors using PCA-based feature selector and correlation-based feature selector. An attack detection method was presented that assists in identifying various attack scenarios utilizing a statistical method. The statistical approach towards the network analysis based on deviation from the standard conduct of the network traffic is employed.

Rejimol and Ciza [9] proposed to decrease false positive and false negative errors, boosting precision, recall and keeping detection accuracy by employed supervised ML algorithms to evaluation and ranking of some features. This work suggests the effectiveness of performance-based classifiers particularly the performance algorithm of Ad boost with Random Forest as the base classifier. Research shows that the use of SDN brings a new opportunity to defeat denial of service attacks in an electronic computing environment.

Gnanapriya et. al.[10] worked to develop the work of IDS by identifying the reduced features Information. Gain, and the Gain Ratio were adopted as techniques for features selection. Additionally, they utilized the Feature Vitality Based Reduction Method (FVBRM method) for getting and distinguishing the reduced group of important features. They used a Greedy Forward Selection method for feature selection and selected 20 attributes out of 41 features. The classification accuracy of the KNN algorithm was close to 100%.

Opeyemi et.al. [11] suggested a method named an "ensemble-based multi-filter feature a selection of EMFFS" which brings together the output of four filtering methods to select features to achieve and improve the selection of features. Their results show accuracy of 99.67%

Suleman et. al. [12] focused on identifying some of the most likely attributes of an attack denial service based on the calculation of weights using entropy calculations. The proposed method relies mainly on using Shannon entropy and granular to calculate the weights to select the most important features of their impact in the denial of service attack.

III. DDoS ATTACKS

One of the well-known threats to cyber security is Distributed Denial-of-Service (DDoS) attacks in which the victim network is bombed by the whole of the worthy number of attack packets originating from many computers. The hosts that operate these tools have an attack that they call zombies and can be attacked by an attacker taking something. A successful DDoS attack manages to degrade the subject and cause serious harm to critical infrastructure elements. A Denial-of-Service DoS attack is a malicious try to consume a service and thus denying reach to it for its legitimate users. A Distributed Denial-of-Service (DDoS) the attack is a subset of the DoS attack, which as the determine suggests, is carried out in a distributed fashion. To obtain this, attackers commonly employ what is called a botnet to execute the attack [13] .To execute the DDoS attack, the attacker determines to discover the vulnerable hosts which can be utilized to start the DDoS attack by applying variety of survey methods named random scan, multiple scan list, topographic scan, local subnet scan, scan switch, and partitioned permutation scanning. After knowing the vulnerable devices, the attacker injects malware or attack tools into the device by utilization the weakness in the machines [14] .The Figure I below illustrates the DDoS attack structure.

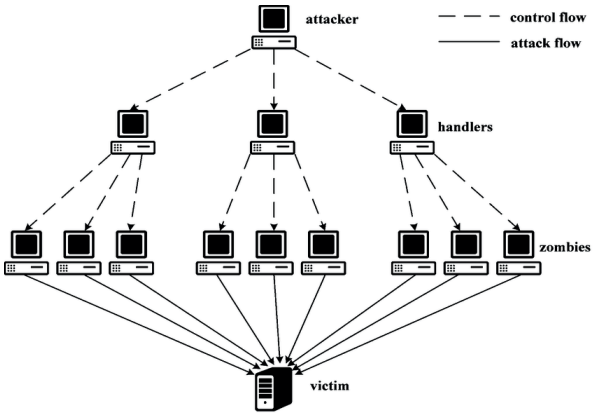


Figure 1: Structure of DDoS attacks

IV. DESIGN AND METHODOLOGY

The purpose of the suggested method is to reduce feature sets considerably while maintaining and improving classification accuracy by using back propagation neural networks. In suggested method, we combine the method of feature selection and dimension reduction called (FSDR) for selecting important features that influence in DDoS attacks. In literature studies, it has been found that most feature selection and dimensionality reduction methods are biased towards considering attributes with many different values. Despite, attributes with quite small information value to seem to receive unfair choice [15]. We use BPNN algorithm to evaluate the approach for detecting the DDoS attack to diminish the false positive value. The proposed method includes three stages. Figure II shows the entire system flow of the proposed model. Proposed framework consists of following stages:

⇒ Phase I (Data Collection and Pre-processing)
- Selecting NSL-KDD dataset or Alkasasbeh dataset.
- Dataset Preparation, Labeling Dataset, and normalization.

⇒ Phase II (Our algorithm FSDR for feature selection)
- Apply the mutual information for feature scoring.
- Apply Normalization on new feature scoring by empirical cumulative distribution functions to expand variance of feature scoring.
- Determine threshold depend on uncertainty value to select subset of features.

⇒ Phase II (Dimension Reduction)
- Apply developed SVD with new features space.
- Apply PCA for comparison
- Apply SVD for comparison
- Apply datasets without using dimension reduction methods.

⇒ Phase IV (Verification of DDoS attacks Detection)
- Training and testing by using NNs Using the Back-propagation Algorithm.
- Evaluating results of both datasets

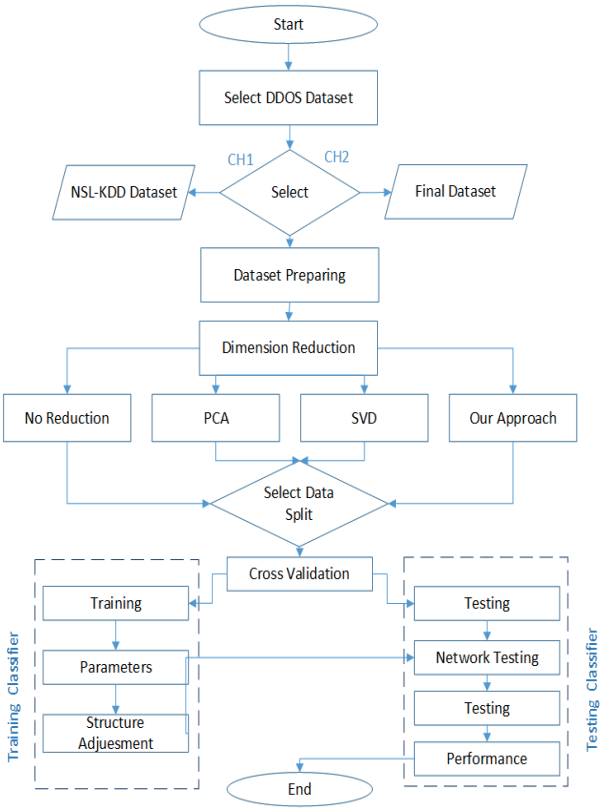


Figure 2: Component architecture of the proposed work

In this search, we used two types of data set. First is NSL-KDD benchmark dataset. The focus of this study on DoS attacks only because it constitutes 78.3% of all attacks. DoS attack to the rest of the attacks occur in a network of six kinds of attacks. The four most important types of DoS attacks were selected out of six as Neptune, Teardrop, Smurf and Back as well as Normal were extracted from the complete NSL KDD data set. The dataset contains 41 features and the selected part of the dataset consists of a training set containing 113051 records and a test set of 15404 records.

Second dataset we used is a new dataset by Alkasassbeh et. al. [1]. They provided a set of data containing modern types of an attack targeting different OSI layers such as Application layer and Network layer. The dataset contains 2160668 records of the normal attack types. We randomly selected 25,000,000 records to test them in our model between normal traffic and four kinds of attack.

A. Phase I: Data Collection and Pre-processing

a. Preprocessing and labeling the dataset

Data preprocessing comprises of a sequence of steps to convert raw data acquired from data extraction into a clean and neat data set before statistical analysis Preprocessing is aimed at evaluating and enhancing the quality of data to enable regular statistical analysis. It includes preprocessing data such as cleaning integration conversion and reduction data[16]. Preprocessing stage, the size of the data indicates that needs to reduce the size and preprocess data to change the data type to labor with the selected classification algorithm. The NSL-KDD and Alkasasbeh datasets as input data sets include several features. However, they are in different formats. Some are in the numeric format, others are in nominal format. Therefore, data sets of these different formats must turn into similar formats and extract from the next stage. The types of nominal features were mentioned in the previous chapter and will be numbered of sequence with the two original data sets. Conducts the conversion processes the following steps:

1. Input NSL-KDD dataset /Alkasasbeh dataset
2. Calculate the probability of (protocol, service, flags, node name from and name to.)
3. Replace nominal value of numeric data
4. Generate a numeric CSV file
5. Read a CSV file and normalize it, using Min-Max normalization
6. Generate a normalized Mat and CSV file.
   Output NSL-KDD Dataset /Alkasasbeh dataset.

b. Normalization

The third stage of data processing, it is performed by a preprocessing module and data for a ready machine learning. Normalization makes it easy to compare different features with a range specified on a different scale, such as (-1 to 1) or (0 to 1). In addition, normalization can accelerate the time needed to train a neural network[17]. We used normalization in three stages of work. First step is the process of converting the nominal features into numeric; we used the min-max equation. Second step is the stage of mutual information for feature scoring using the z-score and at final step is in the pre-classification phase using the data also using z-score equation. The method uses Min-Max normalization strategy between 0 and 1 to normalize.

B. Phase II. Feature Selection

a. Mutual information (MI)

Mutual information MI is an essential connotation in information theory also it is broadly employed in artificial intelligence (AI). MI indicates the presence of the information shared among two variables and measuring the extent to which one variable is known will be reducing the value of uncertainty for the other. Higher mutual information means that the uncertainty is largely reduced, Low mutual information means a small reduced in uncertainty, zeroing mutual information among two random variables intends that the variables are not dependent. The way it works depends on if there is a lot of information shared among the feature and class of labeled data then this is a strong signal that this feature is significant and beneficial in identifying class members from other. The idea about the uncertainty concerning random variables is adopted by its entropy. The concept of entropy included discrete variables and continuous variables.

b. Empirical Cumulative Distribution Functions

In statistical methods, empirical distribution function indicates the distribution function correlated by empirical measurement from the sample. The Empirical Cumulative

Distribution Function ECDF is a valuable assessment of the cumulative distribution function CDF to a random variable. We use ECDF to expand weights variance of features by normalizing it that has obtained form MI and the determent threshold of uncertainty value that formed the points in the sample. The empirical CDF is calculated by ranking each sample (separately for each variable) and then rescaling the integer rank values to the range (0, 1). The ECDF is nearly related to the cumulative frequency and it is determined by getting no assumptions about the underlying distribution. It allocates a probability of 1/n to each point of data n. The order arranges the ascending order of the values, and the sum of the allocated probabilities is calculated up to including each observation in a sample.

C. Phase III: Dimension Reduction

a. Enhanced Singular Value Decomposition (SVD)

SVD is considered to be one of the most important and common methods factorized matrix. SVD provides an accurate description of any matrix, and as well as facilitate the removal of the insignificant parts of this representation to provide an approximate representation of any number of respectable dimensions.

This part is considered the last of the second stage and after completing the selection of subset features from the original datasets by using MI. We will use SVD's enhanced dimension reduction algorithm for the new set of features. Our model depends on the regular SVD but our idea is to rank the Eigenvalue which is in diagonal and accelerated Singular Value Decomposition.

b. PCA algorithm

Principle Component Analysis (PCA) is one of the linear methods that reduce the dimensions of data by analyzing the variance between the factors. PCA investigates the correlation between a couple of attributes to determine that important one and leads to a linear drawing of data into a less dimensionality space so that the contrast of the data is magnified in the low dimensionality space. PCA builds the data correlation matrix and calculates the eigenvector corresponding to the largest eigenvalue component used to reconstruct the majority factors of the variation from the original data.

D. Phase IV: Neural Network DDoS Detector

a. Back-Propagation

This method enables the error derivative calculation between the actual output and the target which is derived independently for all weights in the network. In multi-layered Perceptron NNs, the backpropagation algorithm is utilized to calculate the necessary modifications after random selection of the network weights.

b. Residual Network

Residual Network is a neural network structure that resolves the difficulty of fading gradients in a way that is as simple as possible If you encounter problems sending the gradient signal in the reverse direction let's set a shortcut for each layer in the network and proceed smoothly.

E. Training and Testing Stages

All stages of training and testing in classifier BP-ANN for both data sets NSL-KDD and Alkasasbeh were validated by 5 cross-validation to evaluate the model performance.

The aim of employ cross-validation is to use the whole available data set in the learning stage. To accomplish this cross-validation the data is randomly split into K fold. As every fold, the whole data set is employed, however it is a learning set. It is employed as a test set. The mean error for each fold produces a small biased estimator [18]. We applied 5-fold Cross Validation in the study by using MATLAB. The datasets are randomly split into 5 segments where the layer is represented in roughly the like proportions as in a whole set of data. Each segment is kept in turn and the learning stage is trained on the remaining 4/5 next the error rate is computed on the standing segment. The learning phase was carried out a total of 5 times on separate various training sets and subsequently, the average error rates of five resulted in a general error calculation. The consequence of the classifier BP-NN is the part that really makes us understand whether our DDoS detection model is it operating correctly or not.

V. RESULTS AND DISCUSSION

In this section, we will review the general performance results of the BPNN by using NSL-KDD and Alkasasbeh datasets based on the training and testing. We empirically evaluate Guided Dimension Reduction Features Selection (GDRFS) by comparing it with some of the dimension reduction methods based on two data sets NSL-KDD and Alkasasbeh. We essentially observe the difference between the returned feature subsets related to different dimension reduction methods. Hence, we respectfully present the results of the NN-BP classifier, since the classification performance is the ultimate evaluation of a criterion. As one of the basic methods of the data sets of the proposed model in the selection of features and the reduction of dimensions using MI and developed SVD and their effect to detect DoS and DDoS attacks. We will review the results of four experiments using the SVD algorithm, the PCA algorithm, and the two sets of data without using the reduction methods. These algorithms are used to reduce the space feature of 41 and 27 and the features of both data sets and the classification of four types of DDoS attacks for both data sets. The results were done by using the MATLAB 2014a, program language to implement our proposed approach.

There are four sets of experiments implemented in the study, which are our model GDRFS experiment, PCA experiment, SVD experimental and non-dimension reduction experiment. All experiments are implementing in Neural Network BPNN classifier to testing two datasets NSL-KDD and Alkasasbeh data set for detection DoS and DDoS attacks in both multi-class datasets and the binary class dataset. The performance results of the Neuronal Neuron Network with both the multi-layer and the binary classes were obtained in two sets of datasets with the highest accuracy of 99.1, 97.92, 98.09, and 98.80 with our GDRFS compared to the other three trials, respectively.

The results indicate that the higher classification accuracy is achieved by a diminutive number of features during the suggested method is employed as feature selection and reduced

dimensionality of big dimensional datasets. That is achieved by use of MI with the ECDF estimator is the first guide in selecting the best features of the total data set, where the descending order of the original group values is performed according to the highest exchange information among the variables. The uncertainty threshold value is usually set to a higher value so that the nonnegative value and not very small in the mutuality of information between the two variables above the threshold and can be selected to subset features. Where ECDF determines the uncertainty values to ignore negative and non-useful values based on MI's work, finds an optimal subset of all attributes and removes irrelevant and redundant attributes. Which identified 17 features out of 41 in the first data set and 8 out of 27 for the second data set.

Since objects in the MI feature selection dataset are not ideal for classification, we used another guide, i.e. the improved SVD algorithm to reduce dimensions. SVD only uses 4 features of mapped data, instead of all 17 features and 6 features of the newly selected in early phase, instead of all 8 features, based on reconstruction. This means that maintaining a 96% variance ensures that the information dropped is not much.

VI. CONCLUSIONS

In this study, we introduced an optimal feature selection and dimensionality reduction with a back-propagation neural networks algorithm for detecting DoS and DDoS attacks in two datasets. There are four kinds of DoS attacks (Neptune, Teardrop, Smurf, Back) on the NSL- KDD data set. The other dataset contains modern types of attack DDoS distributed between the layers of Application and Network and contains four types of DDoS attacks (Smurf, UDP-FLOOD, SIDDoS, HTTP-Flood). Some of the feature selection methods have a specific bias in considering attributes having many different values. However, attributes with very low information value seem to receive unfair profits. Therefore, in the feature selection phase, an MI algorithm was used to rank all data sets with scoring for each feature by descending order. In addition, empirical cumulative density function (ECDF) was used for expanding the variance between weights MI and then selecting subset features to depend on the uncertainty value in MI theory. The Negative scoring values are ignored because they are useless and positive values are chosen. This is considered the first guide to selecting a partial set of features in the first stage. On the dimensionality reduction side, there are many ways to minimize dimensions in the dataset such as PCA, SVD, etc. In our research, SVD was adopted because it works by analyzing matrices easily. We enhanced SVD algorithms. This part is considered the second stage and after completing the selection of subset features from the original datasets by using MI. We will use SVD's enhanced dimension reduction algorithm for the new set of features. Our model, which depends on the regular SVD, but our idea is to rank the Eigenvalue which is in diagonal in two cases and accelerated Singular Value Decomposition. Used to reconstruct the major factors of the variation of the original data at ration 96%. To find the most important features, first, NSL-KDD dataset has reduced the dimensions from 41 to 4. In the modern dataset of DDoS attacks created by Alkasasbeh, its features was reduced from 27 to 6. We obtained the highest classification accuracy by the Neural Network BPNN classifier to testing two datasets NSL-KDD and Alkasasbeh dataset for detection DOS and DDoS attacks in both multi-class datasets and the binary class data set. The performance results of the Neuronal Neuron Network with both the multi-layer and the binary classes were obtained in two sets of datasets with the highest accuracy of 99.1%, 97.92%, 98.09%, and 98.80%. We compared the GFSDR method with the PCA and SVD reduction methods as well as the two sets of data tests with the full set of features.

REFERENCES

[1] Alkasassbeh, M , et al , “Detecting Distributed Denial of Service Attacks Using Data Mining Techniques” *International Journal of Advanced Computer Science and Applications*, 2016 7(1)

[2] S Al-Helli, “Guided Feature Selection and Dimensionality Reduction Method for IDS Improvement in DDoS Attacks”, M Sc Dissertation, Dept Comp Eng , Univ of Turkish Aeronautical Association, Ankara, Turkey, 2017

[3] Balkanli, E , A N Zincir-Heywood, and M I Heywood “Feature selection for robust backscatter DDoS detection” *2015 IEEE Local Computer Networks Conference Workshops (LCN Workshops)*

[4] Haddadi, F , et al “Intrusion detection and attack classification using feed-forward neural network” *IEEE Computer and Network Technology (ICCNT), 2010 Second International Conference.*

[5] Han, J , J Pei, and M Kamber, *Data mining: concepts and techniques* 2011 Elsevier

[6] Sasikala, S , S A alias Balamurugan, and S Geetha, “Multi filtration feature selection (MFFS) to improve discriminatory ability in clinical data set” *Applied Computing and Informatics, 2014.*

[7] Liao, Q , et al “Feature extraction and construction of application layer DDoS attack based on user behavior” *IEEE Control Conference (CCC), 2014.*

[8] Buragohain, C , et al , “Anomaly based DDoS Attack Detection ” *International Journal of Computer Applications, 2015. 123(17).*

[9] Robinson, R R and C Thomas “Ranking of machine learning algorithms based on the performance in classifying DDoS attacks” *IEEE Recent Advances in Intelligent Computational Systems (RAICS),. 2015.*

[10] Fadlil, A , I Riadi, and S Aji, “A Novel DDoS Attack Detection Based on Gaussian Naive Bayes” *Bulletin of Electrical Engineering and Informatics, 2017* 6(2): p 140-148

[11] Osanaiye, O , et al “Ensemble-based multi-filter feature selection method for DDoS detection in cloud computing” *EURASIP Journal on Wireless Communications and Networking, 2016* 2016(1): p 130

[12] Khan, S , et al ,”Feature Selection of Denial-of-Service Attacks Using Entropy and Granular Computing” *Arabian Journal for Science and Engineering, 2017.* p 1-10

[13] Mansfield-Devine, S , “The growth and evolution of DDoS” *Network Security, 2015.* 2015(10): p 13-20

[14] Geoff Huston, A , “Network Service Models “, *The Internet Protocol Journal, . June 2013*

[15] Strickland, J , “Predictive modeling and analytics” *2014: Lulu. com*

[16] Malley, B , D Ramazzotti, and J T -y Wu, “Data Pre-processing, in Secondary Analysis of Electronic Health Records” *2016, Springer.* p 115-141

[17] Krizhevsky, A , I Sutskever, and G E Hinton “Imagenet classification with deep convolutional neural network” *Advances in neural information processing systems. 2012.*

[18] Sterlin, P , “Overfitting prevention with cross-validation” *„Supervised Machine Learning Report,* 2007 83

Phasor particle swarm optimization for solving problems of power flow and pricing in the electricity market using the supply chain equilibrium model

M. M. JEVTIĆ[1], M. D. JEVTIĆ[2], J. RADOSAVLJEVIĆ[2], D. KLIMENTA[2] and S. ARSIĆ[1]
[1]Technical Faculty in Bor, University of Belgrade, Bor/Serbia, mjevtic@tfbor.bg.ac rs
[1]Technical Faculty in Bor, University of Belgrade, Bor/Serbia, saarsic@tfbor.bg.ac rs
[2]Faculty of Technical Sciences Kosovska Mitrovica, University of Prishtina in Kosovska Mitrovica, Kosovska Mitrovica/Serbia, miroljub.jevtic@pr.ac rs
[2]Faculty of Technical Sciences Kosovska Mitrovica, University of Prishtina in Kosovska Mitrovica, Kosovska Mitrovica/Serbia, jordan radosavljevic@pr.ac rs
[2]Faculty of Technical Sciences Kosovska Mitrovica, University of Prishtina in Kosovska Mitrovica, Kosovska Mitrovica/Serbia, dardan.klimenta@pr.ac.rs

Abstract **- This paper proposes the application of the recently developed meta-heuristic algorithm named phasor particle swarm optimization (PPSO) for solving the problem of power flow and pricing in the deregulated electricity market. The market is represented by an equilibrium supply chain (SC) model in which the participants of one tier are competitive with each other but at the same time cooperate with the participants of other tiers of the SC. The cost functions in this model are nonlinear and non-separable. The objective function in determining optimal power flows and nodal prices in the SC is a function of total profit of SC. The results of PPSO application are compared with the results of application of numeric method named modified projection method (MPM) which was applied on the same model in the published literature. Also, the PSOS-CGSA, as one of the latest meta-heuristic algorithms, is used in the paper for comparison. The results show that PPSO gives the better results compared to MPM and PSOS-CGSA. Moreover, in the case of PPSO application the equilibrium conditions are fully satisfied while in the case of MPM the equilibrium conditions are satisfied with small error.**

Keywords **- electricity market, meta-heuristics, phasor particle swarm optimization, supply chain**

I. INTRODUCTION

In the deregulated electricity market, electricity trading takes place on different trading paths such as day-ahead markets, future markets and forward contracts. Different supply chain (SC) models have been used in the literature to analyze these markets [1-3]. One of the models used to analyze the electricity market is the equilibrium model of SC [4]. This model is characterized by the condition that the participants of SC of one tier (traders, for example) are competitive with each other, but each of them is associated with participants in the other layer (for example, generators) with which they are not competitive. In the equilibrium model of SC the condition for determining the node prices is that the total amount of goods (electricity) entering the node is the same as the amount of goods leaving the node. The Equilibrium SC model is also characterized by a large number of controls and dependent variables and functions of different costs in the SC. In [3], the problem of determining the optimal power flows and nodal prices in the SC of electricity market was solved, first, by deriving a variational inequality based on the equilibrium state in the SC and, then, by solving this inequality using the numerical mathematical method - the modified projection method (MPM) [4-5]. In this way, optimal power flows and nodal prices in the SC were obtained, whereby the function of the total profit in the SC was taken as an objective function. The number of iterations of this MPM application was been 230 to 633, depending on the numerical example that was solved, and the equilibrium condition in the SC is approximately satisfied.

The aim of this paper is to apply one of the meta-heuristic optimization methods on the mentioned model of SC and get a better solution than the solution obtained using MPM. It is known that meta-heuristic methods are suitable for solving nonlinear complex problems of large dimensions. Therefore, the paper uses one of the recently developed meta-heuristic methods based on the phasor particle swarm optimization (PPSO) algorithm [6-7] and represents an improved, well-known particle swarm optimization (PSO) algorithm [8]. By applying this algorithm, better solutions are obtained in terms of: the amount of total profit, computation time (number of iterations) and satisfying the equilibrium condition in the SC.

II. THE SC MODEL OF ELECTRICITY MARKET

Figure 1 shows the scheme of equilibrium SC in a deregulated electricity market, which is similar to corresponding scheme given in [3]. The scheme contains three tiers with a certain number of nodes in each tier. In the nodes of the same tier there are SC participants who are competitive with each other, but are uncompetitive in relation to the participants of the other tiers. The links between the tiers represent the rights of participants in the SC to sell or to buy the electric energy and they are not physical. Links 1, ..., q, ..., Q represent variants of the transmission lines used to transmit electrical energy (power). In the first tier are generators (g) which total number is G, in the second tier are suppliers (s) which total number is S and in the third tier are demand

markets (d) which total number is D. Generators sell electricity to suppliers, suppliers sell electricity to demand markets. Suppliers pay service to provider, for transmission and distribution of electric energy. In this scheme, suppliers include retailers, power marketers and brokers. The generators sell the electric energy to the suppliers over the day-ahead spot market, future market or forward contracts.

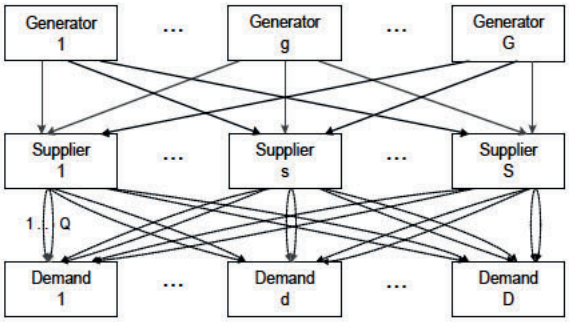


Figure 1: Scheme of the SC for development of a model.

The generator profit (Π_g) can be expressed by the following function:

$$\Pi_g = \sum_{s=1}^{S} p_{gs} P_{gs} - F_g(P_g) - \sum_{s=1}^{S} C_{gs}(P_{gs}) \quad P_{gs} \geq 0, \forall s \tag{1}$$

where indexes g and s denote the generator and supplier, respectively; P_g is the power that is produced from generator g; P_{gs} is the power that is transferred between generator g and supplier s; p_{gs} is the price for transfer of the power from generator g to supplier s in the SC; C_{gs} is the cost of activity at generator g for transfer the power from generator g to supplier s (transaction cost); F_g is the production cost of generator g.

Similarly, the supplier profit (Π_s) can be expressed by the function:

$$\Pi_s = \sum_{d=1}^{D}\sum_{q=1}^{Q} p_{sdq} P_{sdq} - C_s(P_{gs}, P_{sdq}) - \sum_{g=1}^{G} p_{gs} P_{gs} - \\ - \sum_{g=1}^{G} C'_{gs}(P_{gs}) - \sum_{d=1}^{D}\sum_{q=1}^{Q} C_{sdq}(P_{sdq}) \tag{2}$$

where index d denotes the consumer; P_{sdq} is the power that is transferred between the supplier s and consumer d via the transmission system q; C_s is the operating cost of the supplier s; C_{sdq} is the cost of activity at supplier s for transfer of the power from the supplier s to the consumer d via the transmission system q; C'_{gs} is the cost of activity at the supplier s for power transfer from the generator g to the supplier s; p_{sdq} is the price (at supplier s) for the power transmission from the supplier s to the consumer d via the transmission system q; p_d is the price of the power at consumer d.

There are the constraints for each supplier s, which can be expressed as follows:

$$\sum_{d=1}^{D}\sum_{q=1}^{Q} P_{sdq} = \sum_{g=1}^{G} P_{gs} \quad P_{gs} \geq 0, \forall g, \; P_{sdq} \geq 0, \forall s. \tag{3}$$

For equilibrium state of the SC there is a balance of the power, at each demand market d, as follows:

$$P'_d(p_d) = \sum_{s=1}^{S}\sum_{q=1}^{Q} P_{sdq}, \quad p_d > 0 \tag{4}$$

where P_d is the power at the consumer d, that depends on the price p_d at the consumer d. The function $P'_d(p_d)$ is usually obtained by approximating the experimentally demand curve [2]. The second tier price p_{sdq} associated with the power supplier s can be expressed as:

$$p_{sdq} = p_d - c_{sdq}$$

where c_{sdq} is the unit transaction cost between the power supplier s and the consumer d via the transmission system q.

The price p_{gs} is determined from the Nash equilibrium [9] that underlying the non-cooperative behavior of generators when each generator determines his optimal power generation. Based on this concept, the following expression can be derived from (1):

$$p_{gs} = \frac{\partial F_g(P_g)}{\partial P_{gs}} + \frac{\partial C_{gs}(P_{gs})}{\partial P_{gs}} \tag{5}$$

The interpretation of (5) is as follows: the price p_{gs} is equal to the sum of the marginal production cost (first term on the right) plus marginal transaction cost of the generator, associated with that supplier (second term).

The units of variables in this model are: MW (for power flows), $/h (for profits, prices and costs).

III. SOLUTION METHOD

PSO is the most common optimization meta-heuristic algorithm based on swarm population behavior in a nature [8]. In the PSO, swarm units are particles, each of which represents a potential solution of the optimization problem. The particles are presented by position vector and the velocity vector. They gradually change their positions and velocities in the search space moving towards the best solution. The starting positions were chosen randomly within the given limits of possible solutions. In the next iterations, new positions (solutions) are calculated using the new velocities (adaptations when changing solutions) applying the following equations:

$$\mathbf{v}_i(t+1) = w(t)\mathbf{v}_i(t) + C_1 r_1 \left(\mathbf{pbest}_i(t) - \mathbf{x}_i(t)\right) + \\ + C_2 r_2 \left(\mathbf{gbest}(t) - \mathbf{x}_i(t)\right) \tag{6}$$

$$\mathbf{x}_i(t+1) = \mathbf{x}_i(t) + \mathbf{v}_i(t+1) \tag{7}$$

were $w(t)$ is the inertia weight, which makes a balance between local and global search, C_1 and C_2 are the acceleration control coefficients, r_1 and r_2 are the uniformly distributed random numbers within [0,1], $\mathbf{pbest}_i(t)$ is the best position of i^{th} particle (personal best position), and **gbest** is the global best position in the whole population.

PPSO is improved PSO based on modeling the coefficients C_1 and C_2 using the phase angle (θ). Thus, PPSO becomes nonparametric algorithm. In the PPSO the inertia weight $w(t)$ is set to zero. The equations for updating the values of velocity, particle position and phase angle will be the following:

$$\mathbf{v}_i(t+1) = \left|\cos\theta_i(t)\right|^{2\sin\theta_i(t)} \cdot \left(\mathbf{pbest}_i(t) - \mathbf{x}_i(t)\right) + \\ + \left|\sin\theta_i(t)\right|^{2\cos\theta_i(t)} \cdot \left(\mathbf{gbest}(t) - \mathbf{x}_i(t)\right) \quad (8)$$

$$\mathbf{x}_i(t+1) = \mathbf{x}_i(t) + \mathbf{v}_i(t+1) \quad (9)$$

$$\theta_i(t+1) = \theta_i(t) + \left|\cos\theta_i(t) + \sin\theta_i(t)\right| \cdot 2\pi \ . \quad (10)$$

The general structure of the PPSO algorithm is as follows:
Initialization

1. Define the objective function $F(\mathbf{x_i})$ and a space of possible solutions **X.**
2. Generate the number of search agents (N) (possible solutions) which are randomly selected between the limits of the control variables values. Set the iteration counter t=1.

Calculation and iterative procedure

3. Calculation the fitness value for each agent in the current population.
4. Generate the new population of agents, updating the current velocity and current position of agents, using the equations (8) and (9).
   The phase angle θ_i is selected through uniform distribution U (0,2π) and corresponds to the randomly selected x_i. The phase angle is updated using the equation (10).
5. The process of updating the velocities and the positions is stopped by meeting a stop criterion.
6. Report the best solution. Stop.

IV. Application of PPSO for Solving Problems of Power Flow and Prices

Application of PPSO for solving problems of power flow and prices in the electricity market using above equilibrium model of SC can be described in the steps shown in Table 1.

The PPSO is implemented in MATLAB R2017a computing environment and run on a 1.6 GHz, PC with 3.0 GB RAM. The population size and maximum number of iterations are set to 50 and 200, respectively.

Table 1: Steps of PPSO application for solving problems of power flow and prices

Step 1	Read the input data: SC configuration, cost functions of generators, suppliers and transaction, and demand functions.
Step 2	Specify the objective function. Specify the control and depended variables. Specify the limits of variables.
Step 3	Define the number of search agents (population size). Set the maximum number of iterations. Generate the initial random search agents.
Step 4	Calculate the values of the objective function (fitness
Step 5	Apply the PPSO operators (8), (9), (10) to create a new population of agents, which are improved solutions of problem.
Step 6	Repeat steps 4-5 until the stop criterion is meet, i.e. the maximum number of iterations.
Step 7	Report the optimal solutions.

V. Simulation Results

The PPSO has been tested on the SC numerical example which was used in [3,10], for testing the MPM. The power generating, transaction, and operating cost functions, and the demand functions of the numerical example are given in Table 2. The SC numerical example includes 3 generators, 2 suppliers and 3 demand markets. The objective is to maximize the total profit of SC ($\Pi_g + \Pi_s$). The obtained optimal values of power transacted in the SC, demand prices and prices for transfer of the power from generator g to supplier s are listed in Table 3.

Figure 2 shows the convergence profile of PPSO and PSOS-CGSA in solving this problem. The PSOS-CGSA is one of latest algorithms which was developed in [11] and is used in this paper for comparison of results. It may be observed from Figure 2 that PPSO and PSOS-CGSA converge to their global optimums in 20 and 55 iterations, respectively. The MPM, which was applied in [3,10], converged in 232 iterations for same test example.

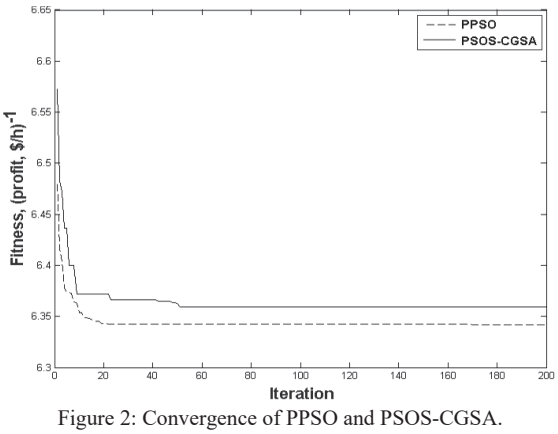


Figure 2: Convergence of PPSO and PSOS-CGSA.

The results shown for the PPSO and the PSOS-CGSA are the best values obtained over 30 consecutive runs. Table 4 shows the comparison of minimum, maximum, and mean

values and standard deviations of the results obtained by PPSO and by PSOS-CGSA over 30 runs. It can be seen that the proposed PPSO provides better solutions than PSOS-CGSA.

From Table 3 follows that the MPM applied in [3,10], resulted in lesser values of maximum profit than the PPSO and PSOS-CGSA which are applied in this work. However, the error of equilibrium condition (3) is zero in the case of PPSO and PSOS-CGSA application while that error is $8\cdot10^{-4}$ MW in the case of MPM application. This means that the application of the PPSO provides more accurate results and the higher maximum profit for the lesser number of iterations than the application of the MPM.

Table 2: Numerical example [3, 10], for simulation

Power generating cost functions, $F_g(P_g)$ $F_1(P_1,P_2)=2.5P_1^2+P_1P_2+2P_1;\ F_2(P_1,P_2)=2.5P_2^2+P_1P_2+2P_2$ $F_3(P_1,P_3)=0.5P_3^2+0.5P_1P_3+2P_3$
Transaction cost functions faced by the power generators and associated with transacting with the power suppliers, $C_{gs}(P_{gs})$ $C_{11}(P_{11})=0.5P_{11}^2+3.5P_{11};\ C_{12}(P_{12})=0.5P_{12}^2+3.5P_{12}$ $C_{21}(P_{21})=0.5P_{21}^2+3.5P_{21};\ C_{22}(P_{22})=0.5P_{22}^2+3.5P_{22}$ $C_{31}(P_{31})=0.5P_{31}^2+2P_{31};\ C_{32}(P_{32})=0.5P_{32}^2+2P_{32}$
Operating cost functions of the suppliers, $C_s(P_{gs})$ $C_1(P_{g1})=0.5(\sum_{g=1}^{3}P_{g1})^2;\ C_2(P_2)=0.5(\sum_{g=1}^{3}P_{g2})^2$
Unit transaction costs between the power suppliers and the consumers, c_{sdq} $c_{111}=q_{111}+5,\ c_{121}=q_{121}+5;\ c_{131}=q_{131}+5,\ c_{211}=q_{211}+5$ $c_{221}=q_{221}+5,\ c_{231}=q_{231}+5$
Power of demand, P_d', as a function of the price p_d at demand market: $P_1(p_1,p_2)=-2p_1-1.5p_2+1100;P_2(p_2,p_2)=-2p_2-1.5p_1+1100$ $P_3(p_1,p_3)=-2p_3-1.5p_1+1200$

Table 3: Comparison of simulation results

Control variable	PPSO	PSOS-CGSA	MPM [3]
P_{11} (MW)	12.50182	12.54946	14.2762
P_{12} (MW)	12.50182	12.54946	14.2762
P_{21} (MW)	12.50182	12.54946	14.2762
P_{22} (MW)	12.50182	12.54946	14.2762
P_{31} (MW)	59.99781	59.91425	57.6051
P_{32} (MW)	59.99781	59.91425	57.6051
P_{111}(MW)	20.00073	20.00567	20.3861
P_{121}(MW)	20.00073	20.00567	20.3861
P_{211}(MW)	20.00073	20.00567	20.3861
P_{221}(MW)	20.00073	20.00567	20.3861
P_{131}(MW)	45.00000	45.00185	45.3861
P_{231}(MW)	45.00000	45.00185	45.3861
p_1 ($/h)	302.85673	302.85391	302.6367
p_2 ($/h)	302.85673	302.85391	302.6367
p_3 ($/h)	327.85746	327.85772	327.6367
Maximal profit, ($/h)	15767.81557	15760.59073	15218.6668
Error of equilibrium condition (3), (MW)	0	0	$8\cdot10^{-4}$

Table 4: Statistical evaluation of the simulation results

	min	max	mean	Standard deviation
PPSO	15747.5355	15767.8156	15765.6592	$1.7982\cdot10^{-8}$
PSOS-CGSA	15626.9534	15760.5907	15738.3654	$1.0426\cdot10^{-7}$

VI. Conclusions

This paper proposes the application of the recently developed meta-heuristic optimization algorithm named PPSO for the computation of power flow and prices in the electricity market presented by the SC equilibrium model. Functions that describe the costs of participants in the SC (generators, suppliers, and transmitters) are nonlinear and non-separable, which makes it difficult to apply mathematical optimization methods. Participants in one layer of SC are competitive with each other while they are cooperative with participants in other layers. In this paper, we tested PPSO on a numerical example of electricity equilibrium SC and compared the results with the results of the numerical method, MPM, used in the published literature, and with the results of PSOS-CGSA, a meta-heuristic algorithm that was also recently developed. We found that PPSO gives the best results in terms of: the size of the maximum profit in the SC, the computation time (convergence speed), and the accuracy of the results of the computation of power flow and prices in the SC. The PSOS-CGSA gives better results too, compared to the MPM method. The results of this paper can be used for modeling the SC and to handle the behavior of the decision-makers in the electricity market.

Acknowledgment

This work is funded by the Ministry of Education, Science and Technological Development of the Republic of Serbia.

References

[1] F Oliviera, C Ruiz, "Analysis of Futures and Spot Electricity Markets under Risk Aversion," *European Journal of Operational Research*, In Press, Journal Pre-proof, Available online 12 October 2020

[2] F Oliviera, C Ruiz, A Conejo, "Contract design and supply chain coordination in the electricity industry," *European Journal of Operational Research*, vol 227,pp 527-537, January 2013

[3] A Nagurney, D Matsypura, "A Supply Chain Network Perspective for Electric Power Generation, Supply, Transmission, and Consumption," in E J Kontoghiorghes, C Gatu (eds) *Optimisation, Econometric and Financial Analysis,* pp 3-27 Berlin: Springer, 2006

[4] A Bnouhachem, "A modified projection method for a common solution of a system of variational inequalities, a split equilibrium problem and a hierarchical fixed-point problem," *Fixed Point Theory and Applications*, vol 22,pp 1-25, January 2014

[5] G M Korpelevich, "The Extragradient Method for Finding Saddle Points and Other Problems," *Matekon*, vol 13, pp 35-49, 1977

[6] M Ghasemi, E Akbari, A Rahimnejad, S E Razavi, S Ghavidel, L Li, "Phasor particle swarm optimization: a simple and efficient variant of PSO," *Soft Computing*, vol 23,pp 9701-9718, October 2019

[7] Z Ullah, S Wang, J Radosavljević, „A novel method based on PPSO for optimal placement and sizing distribution generation,"

IEEJ Transactions on Electrical and Electronic Engineering, vol 14,pp 1754-1763, September 2019

[8] J Kennedy, R Eberhart, "Particle Swarm Optimization," in Conf Rec *1995IEEE International Conference on Neural Networks,* pp 1942–1948

[9] J F Nash, "Noncooperative games," *Annals of Mathematics*, vol 54,pp 286-298 1951

[10] A Nagurney, *Network Economics: A variational inequality approach. Revised Second Edition*, Springer Science & Business Media B V, vol 10, 2013

[11] Z Ullah, M R Elkadeem, S Wang, J Radosavljević, "A Novel PSOS-CGSA Method for State Estimation in Unbalanced DG-Integrated Distribution Systems," *IEEE Access*, vol 8, pp 113219-113229, June 2020

Investigation of Neurological Rehabilitation Exercise System with Virtual Reality Method

Mesut ÖZARSLAN[1] and Hasan SERDAR[2]

[1] Necmettin Erbakan University, Konya/Turkey, mesutozarslan1@gmail.com
[2] Necmettin Erbakan University, Konya/Turkey, hserdar@erbakan.edu.tr

Abstract **- As the average age of the world population increases, the number of people in need of rehabilitation treatment is also increasing. People with neurological diseases such as stroke, Parkinson's and Cerebral have difficulty performing daily activities such as walking, running and grasping. These neurological diseases are tried to be treated with various rehabilitation techniques. Some of these are traditional treatment methods. In the researches; it has been observed that these methods generally fail to restore the motor functions of the patient. With Virtual reality applications, that have been used in recent years, it has led to the development of alternative treatment methods in the field of rehabilitation. With an exercise program that will provide continuity in physical therapy and rehabilitation, it can be predicted that the motor and functional development of the patients can accelerate the treatment process. In this way, moving the rehabilitation to non-hospital environments such as home environment will give the opportunity to reach more patients at a lower cost. In this study, the effect of virtual reality applications on neurological rehabilitation exercise systems were examined. These are: Smart gloves application for hand rehabilitation systems, Measurement and Exercise Systems using Kinect Sensors for Neurological Rehabilitation such as various studies about neurological rehabilitation exercise system with virtual reality method were examined and collected in a single article. In the content of this study, basic information and important features of these neurological rehabilitation exercise system are mentioned. In the last part of the study, the results of some studies on these neurological rehabilitation exercise systems were given. As a result of the studies, the virtual rehabilitation system can be applied more easily than other applications and at the same time, it allows patients to work at home more comfortably from other neurological rehabilitation exercise system as a performance. In addition, the stands out with its trained models. This article is hoped to help designers who will design neurological rehabilitation exercise system applications.**

Keywords **- Virtual Reality, Neurological Rehabilitation, Exercise System, Physical Therapy.**

I. Introduction

With the increase in the average age of the world population, the number of patients in need of rehabilitation treatment also increases. People with neurological diseases such as stroke, Parkinson's and Cerebral have difficulty performing daily activities such as walking, running and grasping. These neurological diseases are tried to be treated with various rehabilitation techniques that are constantly being developed. Some of the commonly used treatment methods are traditional treatment methods. These methods are generally insufficient to restore the patient's motor functions. Being developed as an alternative; Virtual reality methods such as smart glove applications for hand rehabilitation systems, Measurement and Exercise Systems using Kinect Sensors for Neurological Rehabilitation, various studies related to neurological rehabilitation exercise system have led to the use of new methods in the field of rehabilitation.

These exercise programs, which provide continuity and continuity in physical therapy and rehabilitation, will contribute positively to the treatment process of motor and functional development in patients. Moving rehabilitation to non-hospital environments such as home environment will also provide the opportunity to reach more patients at a lower cost.

In this study, neurological rehabilitation exercise system with virtual reality method was investigated.

II. Methods Used in Neurological Rehabilitation

A. Smart glove applications for hand rehabilitation systems

In rehabilitation treatment, the treatment and rehabilitation of the hands and arms, which are the basic elements of our daily life activities. There are systems using many different technologies and methods developed for hand and arm rehabilitation. In recent years, many wearable rehabilitation systems have been developed. Interest is increasing due to its ease of use and portability. When these applications are combined with virtual reality applications, stimulating the patient's nerves as much as possible during the treatment plays an important role in the healing process.

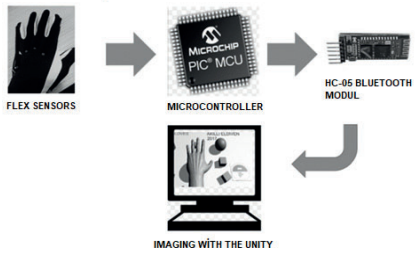


Figure 1: Block diagram of smart glove-based rehabilitation system [2]

With the help of flex sensors placed on the glove, the bending degrees of the fingers are passed through a pre-processing circuit and applied to the analog input of the microcontroller. After the voltage information is converted into angle information with the software encoded in the microcontroller, it is transferred to the computer wirelessly with the help of the bluetooth module. As shown in Figure 2, the hand simulation designed in Unity3D environment simultaneously shows the movements of the person using the glove on the application.

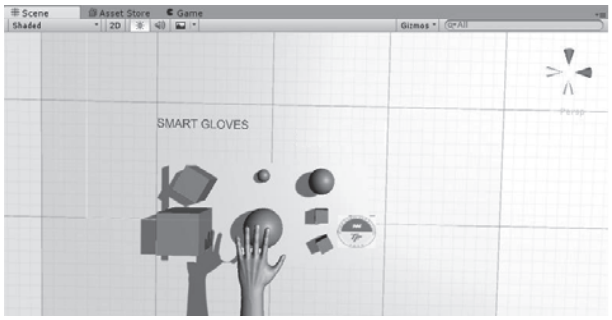


Figure 2: Designed for Hand Rehabilitation system an image from the simulation [2]

B. Measurement and Exercise Systems using Kinect Sensors for Neurological Rehabilitation Systems

When the relevant literature studies were examined, it was found that in individuals treated using the Kinect device, improvement was observed in various diseases such as upper extremity use, balance and mobility in a short time.

In a group study conducted in the United Kingdom, it was stated by rehabilitation specialists that the most common assistive technology for stroke patients is computer games. According to a survey of therapists and physical therapists at a rehabilitation center in Canada, the Nintendo Wii was helpful in treating 76% of inpatients. In addition, 73% of outpatients were reported to adhere to the treatment programs applied with Nintendo Wii.

a) Kinect Hardware Technology

Kinect sensor, which is designed to detect body movements and be used instead of a control device in gaming applications, was developed by Microsoft in 2010. Although Kinect Sensor (Figure-3) came to the market as a part of game consoles, it has become used in different fields such as biomedical, robotics, military, education and security systems. Kinect sensor consists of 4 basic parts: three-dimensional (3D) depth sensor, RGB camera, microphone group and motor.

Figure 3: Kinect Sensor [7]

b) Kinect's Skeleton Detection And Tracking System

With a depth sensor consisting of a black and white CMOS sensor combined with a projector, it emits infrared light to the person opposite it from a distance of 1.5-2.5 meters, and a three-dimensional image consisting of points (Figure-4) is obtained. It is a device that Microsoft launched in 2010, which allows the movements of 30 different joints as hand, arm and leg to be converted into digital data. It also has the ability to perceive sound detection, heart rate, facial expression, facial expressions, and the weight on each limb.

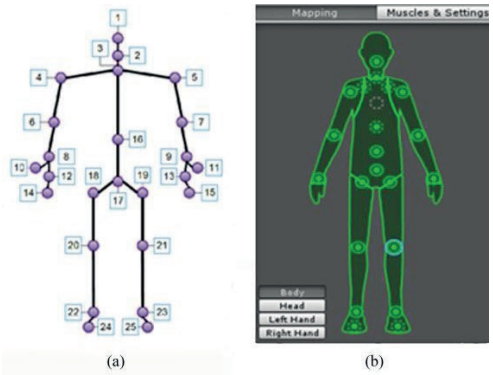


Figure-4: (a) Points captured by the Kinect sensor, (b) Avatar created with the help of dots.[5]

c) Kinect Software Technology

Kinect software can work in conjunction with most software. There are many software libraries that allow the processing of data provided by the hardware. The picture, depth and sound data collected by the sensors are sent to the application after being processed in the "Natural User Interface Library" and this data flow is repeated with the reverse flow.

Figure 5: Kinect application's software and hardware interaction [7]

Unity is a cross-platform game engine developed by Unity Technologies, primarily used to develop video games and simulations for computers, consoles and mobile devices. First announced exclusively for OS X at Apple's Worldwide Developers Conference in 2005, it has since expanded to target 27 platforms.

Six major versions of Unity have been released. At the 2006 WWDC show, Apple ranked Unity first in the Best Mac OS X Graphics Use category.

Examples of games written with Unity engine are Battlestar Galactica Online, Firewatch, Unturned, Layers of Fear, Armello, The Long Dark, and Traffic Racer. Also, Pillars of Eternity and Tyranny, which use the modified version of Unity, are role-playing games made using this engine.[7]

III. RESULTS

A total of 42 patients, including 21 virtual reality groups and 21 control groups, was examined. A total of 35 patients, 19 patients in the virtual reality group and 16 patients in the control group, completed the study and the statistical analysis was made on these data.

Fugl Meyer (FM) scale mean value before treatment; It was 24.32 ± 7.87 (10-36) in the virtual reality group and 19.88 ± 3.79 (12-25) in the control group. The median values were 24 and 20.5 before treatment, respectively. Average FM scale in post-treatment evaluation; 43.05 ± 12.59 (20-60) in the virtual reality group, 34.44 ± 10.53 (18-59) in the control group; median values were determined as 45 and 31.5. In the intergroup evaluation, there was a difference between the pre-treatment FM values of the two groups and it was found to be high in the virtual reality group, there was a difference between the post-treatment FM values of the two groups, and it was also high in the virtual reality group ($p = 0.04$ *, $p = 0.04$ *). When the mean FM scores of both groups before and after treatment were evaluated, a statistically significant improvement was found in both groups ($p < 0.001$*, $p < 0.001$*). FM upper extremity score changes of the groups before and after treatment and when the FM gains between the two groups were compared, no statistically significant difference was found ($p = 0.057$), FM gains are given in Table 1. [6]

	VR GROUP N:19	CONTROL GROUP N:16	P
(Age (year)	69.42 ± 8.55	63.44 ± 15.73	0.19
Duration of illness	88.32 ± 56.32	68.63 ± 39.20	0.25
Gender			0.43
Woman(%)	7 (%36.8)	8 (%50)	
Man (%)	12 (%63.2)	8 (%50)	
Dominant el			0.85
Right(%)	17 (%89.5)	14 (%87.5)	
Left (%)	2 (%10.5)	2 (%12.5)	
Practical hand			0.47
Right(%)	5 (%26.3)	6 (%37.5)	
Left (%)	14 (%73.7)	10 (%62.5)	
SVO Type			0.78
Ischemic(%)	15 (%78.9)	12 (%75)	
Hemorrhagic(%)	4 (%21.1)	4 (%25)	
Fugl-Meyer Upper Extremity			
Before treatment (avr± std. d)	24.32 ± 7.87	19.88 ± 3.79	0,04
After treatment (avr± std. d)	43.05 ± 12.59	34.44 ± 10.53	0,04
Earning points (avr± std. d)	18,74 ± 7,67	13,94 ± 6,58	0,057

Table 1: Demographic characteristics of patients in the virtual reality and control groups and Fugl-Meyer upper limb score gains between groups [6]

In another study; Of the 23 patients included in the study, 11 patients took part in the virtual reality-based exercise (VR) group and 12 patients in the task-oriented exercise (TO) group and completed the study by participating in all sessions.

When the groups were compared in terms of gender, stroke type and involved hemisphere distribution with the "Fisher" exact chi-square test, it was observed that the groups were similar according to these parameters (Table 2).

		Group VR n=11	Group OT n=12	p
Gender	Woman	5 6	3 9	0,304
Stroke Type	Ischemic Hemorrhagic	5 6	7 5	0,537
Involved Hemisphere	Right Hemisphere Left Hemisphere	4 7	5 7	0,795
Age (year)		50,82 ± 15,18	49,00 ± 12,84	0,761
Duration of illness		9,18 ± 5,55	10,67 ± 6,29	0,554
MMDM SCORE		27,91 ± 1,64	28,38 ± 0,71	0,380
EGE-AT SCORE		31,55 ± 14,13	31,42 ± 13,68	0,983
"Line bicestion test" the amount of deviation		0,68 ± 2,56	1,46 ± 1,78	0,414
Wrist joint extension		43,19 ± 14,19	46,25 ± 10,25	0,563
Metacarpophalangel joint extension(*)		25, 00 ± 3,87	24,17 ± 4,17	0,625
Proximal interphalangel joint extension		18,18 ± 3,37	19,17 ± 2,89	0,462
Key grip strength(kg)		1,41 ± 1,07	1,29 ± 1,13	0,801
Fingertip grip strength(kg)		1,00 ± 0,67	0,83 ± 0,72	0,571
Triple grip strength(kg)		*Med. (Min.-Mak.)* 1,0 (0,5 – 2,0)	*Med. (Min.-Mak.)* 0,5 (0,0 – 2,0)	0,053
Hand grip strength(kg)		11,73 ± 4,43	9,50 ± 3,75	0,210

Table 2: Comparison of the range of motion in the initial extension direction and the initial pinchmeter and dynamometer measurement results in terms of gender, stroke type and involved hemisphere distribution [3]

When the experiments examined above were compared, it was observed that virtual reality applications had a positive effect on the treatment process compared to traditional treatment in both studies.

IV. CONCLUSIONS

With the motion sensing system and play, it can be predicted that by helping rehabilitation in various areas from severe chronic pain to stroke, it can activate the plasticity ability of the brain and regain or maintain lost brain functions. In this way, a system can be developed in which patients in need of neurological rehabilitation can exercise in their own living spaces and the treatment can be followed.

Finger bending angles in smart glove applications for hand rehabilitation systems can be measured with the help of flexible sensors. Simultaneous movement of the hand in a simulation program can be achieved wirelessly with these measured values. It was concluded that the smart glove application can be used easily for hand rehabilitation systems and that it will provide visual feedback to the patient who will be treated, and it can make rehabilitation movements more fun without getting bored.

Although there is no obstacle for stroke patients to perform these movements mechanically, the signals sent from the brain cannot be transmitted to the organs sufficiently, making movements such as walking, speaking and grasping become difficult or even impossible. The success of the treatment depends on the patient's ability to continue the movements he made during the treatment given to him in the clinic, in his own living space. It is predicted that the virtual reality system can accelerate the treatment of patients by creating a realistic environment with hardware and software.

V. REFERENCES

[1] Fidan, U ve Özkan, N 2018, "Nörolojik Rehabilitasyon İçin Kinect Sensörlü Ölçüm ve Egzersiz Sisteminin Tasarımı ve Gerçekleştirilmesi", Afyon Kocatepe Üniversitesi Fen ve Mühendislik Bilimleri Dergisi 18, 727-733

[2] Hakan A , Abdullah B ve Ulvi B 2017, "El Rehabilitasyon Sistemleri için Akıllı Eldiven Uygulaması", Marmara Üniversitesi, Bilgisayar Mühendisliği Bölümü, İstanbul

[3] Dereli, E-E , 2010 "İnme Geçiren Hastaların Üst Ekstremite Rehabilitasyonunda Sanal Gerçeklik Temelli ve Göreve Yönlendirilmiş Egzersizlerin El Fonksiyonlarına Etkisinin Karşılaştırılması" Doktora Tezi, İstanbul Üniversitesi, Sağlık Bilimleri Enstitüsü, İstanbul

[4] Kale, B , 2010 "Oyununuzu Kumandasız Oynayın" Tübitak-Bilim ve Teknik Dergisi, 511, 83-84

[5] Shi, Y , & Peng, Q 2018 A VR-Based User Interface For The Upper Limb Rehabilitation Procedia CIRP, 78, 115

[6] Mirzayev, İ , 2015 "İnme Hastalarında Sanal Gerçeklik Eğitiminin Üst Ekstremite Fonksiyonlarına Etkisinin Araştırılması, Uzmanlık Tezi, Başkent Üniversitesi, Fiziksel Tıp ve Rehabilitasyon Anabilim Dalı, Ankara

[7] Evren Ö , 2017 "Sanal Fizik Tedavi Platformu", Yüksek Lisans Tezi, Fırat Üniversitesi, Bilgisayar Mühendisliği Bölümü, Elazığ

A Novel Permutation Based Solution Representation Technique for Vehicle Routing Problems on GPUs

Erdener ÖZÇETİN[1,3], Gürkan ÖZTÜRK[2,3]

1. Department of Industrial Engineering
Hitit University, Corum, Turkey
erdenerozcetin@hitit.edu.tr

2. Department of Industrial Engineering
Eskisehir Technical University, Eskisehir, Turkey
gurkan.o@eskisehir.edu.tr

3. Computational Intelligence and Optimization Laboratory (CIOL), Eskisehir, Turkey

Abstract

In this study, the vehicle routing problem (VRP) which is a well-known NP-hard combinatorial optimization problem is handled on graphic processing units (GPUs). Solving any kind of VRP is extremely hard when the instance size is large. For this reason, researchers tend to solve the VRP with meta-heuristics. Although, many well-designed meta-heuristics produce near-optimal solutions in reasonable time, still a challenge to solve large scale instances. To accomplish this issue, researchers need novel, fast and wisely designed parallel operators for the proposed algorithms. Furthermore, the success of these operators directly depends on the way the solution is represented. This paper offers a new permutation based solution representation technique (π^+) for vehicle routing problems on GPUs. Results show that proposed technique can be used many algorithms to accelerate computations.

1. Introduction

The vehicle routing problem (VRP) can be described as the problem of designing optimal delivery or collection routes from one or several depots to a number of geographically scattered cities or customers, subject to side constraints [1]. Since the VRP introduced by Dantzig and Ramser [2], many researchers have studied many variants of problem. This can be said that popularity of the VRP still ongoing since many applications of problem can be seen in real-life. Besides popularity, it is still a challenge to solve VRP optimally or near-optimally when the instance size increases.

Metaheuristic algorithms are generally cited as tools for solving certain types of optimization problems, such as combinatorial optimization problems like VRP. These algorithms do not guarantee the exact solution developed for optimization problems; but they are techniques that give satisfactory solutions in a short time. They are often used in cases where exact solution algorithms do not provide solutions in a reasonable time.

Whether it is exact method or meta-heuristic still requires intelligent and fast computational techniques. Especially, when the researchers study with very large scale problems, time is one of the most important issues to generate a good solution.

A graphics processing unit (GPU) is a processor designed to rapidly manipulate and alter memory to accelerate the visualization of images. As well as visualization, GPUs are increasingly used in computing. With the motivation of shortening the solution times of optimization problems, it is becoming widespread to develop metaheuristic algorithms on GPUs, mostly using by CUDA environment. However, this technology is an obstacle to the direct use of traditional programming knowledge and skills. In order to develop applications using this technology, it is necessary to understand the architecture of GPU very well.

According to the structure of the problem, the form of representation, the constraints and the developed algorithm must be suitable for simultaneous computation on the GPU. In order to develop the algorithms on GPUs effectively, it is important to define proper representation in accordance with the problem structure which allows to handle constraints of the problem.

In this study, we focus on solution representation of the VRP. We propose a new permutation based solution representation technique which we called π^+. The main contributions of π^+ can be summarized as follows:

- It is easy to manipulate.
- It offers a tremendous speed up opportunity on GPUs
- It can be adaptable for other combinatorial optimization problems.
-

The remainder of this paper is organized as follows. In Section 2, we review the VRP and meta-heuristic algorithms on GPUs. In Section 3, we describe the proposed permutation based solution representation technique and in Section 4 we show the computational experiments. In conclusion, we discuss the efficiency of the proposed technique and the potential for future studies.

2. VRP and Meta-heuristics on GPUs

Researchers commonly propose meta-heuristics algorithms for any variant of VRP. These algorithms can be single solution based such as tabu search, simulated annealing and variable neighborhood search or population based such as genetic algorithm and ant colony optimization. In many studies, these techniques can be hybridized and enriched with local search strategies to give both intensification and diversification properties to the algorithm.

2.1. VRP

There are many types of the VRP in the literature such as capacity constrained, time windowed and open. When the VRP is mentioned, the first thing that comes to mind is the capacity constrained VRP. A vehicle fleet with the same capacity in the capacity-limited VRP moves from a warehouse and serves customers whose demands are known in advance. Each customer gets service from only one vehicle and each

customer's demand is met. The vehicles cannot exceed their capacity and return to the warehouse at the end of service. It is aimed to find the least costly routes under these constraints. In one variant of the problem, there is the additional constraint that the travel time of each vehicle should not exceed a given travel time limit. In open VRP, each route starts at the depot and ends at a customer, visiting number of customers, each once, without returning to the depot.

VRP diversifies according to the needs in application and gains new variants every day. Both technological innovations and the needs of the industry have led to the emergence of many different types of VRP. However, the need to make effective decisions in a short time has led researchers to various studies on VRP's solution methods.

The diversity of models and solution approaches in vehicle routing is enormous (see, e.g., [3] and [4]). For detailed information for variants of problem readers can see [5]. Some meta-heuristic approaches are commonly studied to solve the VRP like tabu search [6], variable neighborhood search [7] and ant colony optimization [8].

2.2. Meta-heuristics on GPUs

Although metaheuristic algorithms search for solution spaces with highly advanced search techniques, the increase in the size of the instances can cause algorithms to not converge to the best solution. Therefore, parallelization is a very important tool that allows the development of new methods for solving large problems.

Parallel computing has been an inevitable way of solving difficult large-scale problems for years. The design and coding of parallel metaheuristic algorithms are directly affected by the platform where the calculation will be made. GPU, which is used as a visualization tool in computer systems, is one of these platforms.

Although metaheuristic algorithms significantly reduce the computational complexity and perform the search for the best solution, the increase in the size of the problems causes the solution times to increase and the solution quality to decrease. In combinatorial optimization problems, the increase in the instance size causes the solution space to grow exponentially. However, computing the objective function and many other calculations require more processing. For these reasons, a parallel design to metaheuristic algorithms comes to the fore. In a broader sense, parallel and distributed computing, the use of metaheuristic algorithms in the design and coding, aims to:

- Speeding up the search,
- Improving the quality of the solution obtained,
- Increasing consistency,
- Solving large scale problems.

With these purposes, parallelization in metaheuristic algorithms occurs in three main dimensions;

- Solution level parallelization: This type of parallelization involves the simultaneous evaluation of the evaluation of a single solution by breaking it into pieces.
- Parallelization at the level of iteration: It means that multiple solutions can be handled simultaneously or that the neighborhood comparison on a solution can be broken down and handled simultaneously.
- Algorithmic level parallelization: It is the simultaneous processing of different metaheuristic algorithms in order to obtain more consistent results.

2.2.1. Literature review

When it is focused to combinatorial optimization problems that handled on GPUs, we can see several studies in the literature.

Tsutsui and Fujimoto [9] presented a parallel genetic algorithm on GPUs for the quadratic assignment problem (QAP) in their study. In the study, 3 to 12 times faster results were obtained than the serial algorithm. In another study [10], the same authors presented a method for the solution of the quadratic assignment problem, consisting of optimizing ant colonies and combining the 2-opt local search algorithm. According to the results, it was emphasized that the solution times on the GPUs are 24.6 times faster than the serial one. Czapinski [11] developed multiple onset tabu search algorithm for the QAP in the CUDA environment in his studies. In the study, some of the literature problems developed for the QAP were addressed and results were obtained in up to 70 times shorter time with solutions that averagely have 1% gaps to optimal solutions. The same problems also studied on GPUs in [12] and [13].

Cecilia et al. [14] developed an ant colonies optimization for the travelling salesman problem and obtained solutions up to 20 times faster. In another study conducted for the same problem, Delevacq et al. [15] achieved an acceleration up to 23.6 by optimizing ant colonies.

Shulz [16] presented a study on applications of local search algorithms for VRP. In the study, the memory operations were emphasized which is one of the bottlenecks when operating on GPUs. In another study for the VRP, Groer et al. [17] have applied local search techniques on GPUs. In another study, Szymon and Dominik [18] developed a parallel banned search algorithm on GPUs for the multi-criteria VRP.

2.2.2. Handling VRP with meta-heuristics

There are some solutions representation methods for VRP, such as permutation based and binary encoding. It is known that representation methods for VRP strictly depends on the constraints of problem. In addition to this, time consuming operations may be done according to representation of problem.

When it comes to parallel algorithms and concurrent computing, representation design of the problem comes to the fore. Davidovic et al. [19] offers the permutation based representation the best for concurrent computation on many types of problems.

3. Proposed Solution Approach

In many CPU-GPU oriented meta-heuristic algorithms for combinatorial initial problems like the VRP are designed as follows: The starting solutions are generated and some complex operations are made on the host. Then the solution is manipulated and sent to device to make parallel computations. These procedure brings many memory operations between host and device which makes GPUs inefficient.

How the solutions will be represented in the GPU is extremely critical in terms of the simultaneous operation of the algorithm. π^+ has been proposed to show the solutions on the GPU, which provides many advantages.

3.1. Representation

The main purpose of the π^+ is to standardize route lengths. We used dummy demand points ($\emptyset$) to ensure that all routes have equal lengths. The distance of all demand points including the warehouse to $\emptyset$ is *0*, and the distance of $\emptyset$ to all demand points is a very large number (*M*). At the same time, the distance of $\emptyset$ to itself is *0*. However, $\emptyset$ does not have demand.

To make the new representation format easy to understand, we will continue to explain π^+ on open VRP. As it is mentioned before, in open VRP, a route starts from depot and ends at a customer. In this example, we assume that all the vehicles are identical. Namely, they have the same capacities. Except for the warehouse and dummy demand points, the demands are arranged in ascending order to obtain the maximum number of different demand points that can be served with one vehicle. The number *l* is obtained by adding *2* to this number. The reason for adding *2* to this value is that, in accordance with open VRP, the first elements in each route are created from the warehouse last elements and the dummy demand points.

The total length of a π^+ is obtained by multiplying the *l* by the number of vehicles. An example of π^+ with 2 vehicles and 8 nodes including the warehouse is as in Figure 1. Accordingly, it is assumed that a maximum of 4 different demands can be loaded on a vehicle according to the vehicle capacity by listing the demands ascending. Therefore, the number *l* is determined as 6. In π^+ every *l* segment represents a route.

0	3	1	4	5	Ø	0	6	2	7	Ø	Ø

Figure 1: An example of π^+ for open VRP

For his example, the objective function of π^+ can be calculated as $\sum_{i=0}^{n-1} cost[\pi^+[i], \pi^+[i+1]]$. Please note that only a dummy demand point is added as the last column and last row in the cost or distance matrix.

3.2. Objective Function Calculation

Objective function calculation is a critical issue for meta-heuristics. For some types of problems may need re-calculation of objective value. π^+ can reduce computation time for many combinatorial optimization problems like the open VRP. This should be considered that each route a block on grid structure on GPU. A vector can be defined on device. Each element of this vector will keep a route's cost.

Random two instances were created to show the effectiveness of the proposed solution representation method in calculating the parallel objective function. According to this, the instance A has 79 customers and a depot with 10 vehicles, and the instance B has 385 customers and a depot with 47 vehicles. We generate random solutions between 2^{10} and 2^{20}. Time of generating these solutions were discarded from comparison. On the other hand, we took into account the time of data transfers to GPU side. In Table 1, the comparisons can be seen in seconds. We can say that the objective function evaluation on GPUs with π^+ for open VRP offers a tremendous speed-up. Of course, it should not be overlooked that algorithms do not just compute an objective function. For this reason, we focused also other operators like local search.

Table 1: Performance comparison on objective function calculation

	A		B	
Number of solutions	**n80k10 CPU**	**n80k10 GPU**	**n386k47 CPU**	**n386k47 GPU**
2^{10}	0.126	0.085	0.356	0.097
2^{11}	0.154	0.088	0.620	0.100
2^{12}	0 300	0.089	1.412	0.104
2^{13}	0 529	0.094	2.500	0.106
2^{14}	1.162	0.096	5.528	0.143
2^{15}	2 285	0.136	10.027	0.152
2^{16}	4 383	0.142	20.497	0.210
2^{17}	8 887	0.151	42.560	0.351
2^{18}	20.770	0.225	84.241	0.676
2^{19}	34.871	0.365	174.168	1.378
2^{20}	73.872	0.649	372.455	3.282

3.3. Local Search Operators

Most of well-designed meta-heuristics use local search operators. We can divide local search algorithms for VRP into intra-route and inter-route. Intra-route algorithms search for some neighborhoods in a route and inter-route algorithms search for neighborhoods between two or more routes.

3.3.1. Intra-route local search operators

The most popular intra-route local search algorithms are Or-opt, 2-opt or 3-opt for the VRP. We implemented 2-opt and Or-opt for open VRP on GPUs to show the efficiency of π^+. Since a similar strategy is followed in all neighborhoods, the operation on GPU is explained in detail only over the 2-opt neighborhood.

Breaking *(i, j)* and *(i + 1, j + 1)* connections by breaking *(i, i + 1)* and *(j, j + 1)* connections with $j \geq i + 2$ in each route, simultaneously on all routes is being tested. The inverse conversion between *(i + 1, j + 1)* performing at the appropriate ones is also performed simultaneously.

While searching in π^+, the warehouse at the beginning of the routes and the $\emptyset$ at the end are excluded from the test. However, in π^+, $\emptyset$ can also take place somewhere else the last point. When looking at the 2-opt neighborhood, the $\emptyset$ in the route is prevented from getting ahead of a demand point, thanks to the

large number of the distance of *Ø* to the demand points in the cost matrix.

This block level parallelization can bring tremendous speed-up to intra-route local search operators.

3.3.2. Inter-route local search operators

Basically, inter-route algorithms make moves between two or more vehicles. When concurrent search between vehicles is concerned, bank conflict should be taken into consideration. In detail, this is not possible to compare a vehicle with more than one vehicle and apply meaningful changes in simultaneous operation. For example, working with an instance consisting of 3 routes, and searches (R=Route) *R1-R2*, *R1-R3* and *R2-R3* on the IOP are best implemented as one-to-one change. In this case, in the comparisons of *R1-R2* and *R1-R3*, an element from *R1* may be replaced by an element from both *R2* and *R3*, and this causes the algorithm to malfunction.

In the proposed parallel method, comparisons up to *integer ((number of vehicles) / 2)* during the search between vehicles are considered simultaneously. To compare all vehicles with each other a loop must be found in the code. In this case, there is a need for an index set that states which routes to compare in each repetition.

Successively named routes in the index set are taken simultaneously in binary comparison. One mutual element change between routes is considered, so that each vehicle will be used at most once. One of the routes cannot be handled more than once in a concurrent transaction. An example of the index set can be seen in Figure 2.

loop						
1	1	2	3	4	5	6
2	1	3	2	4	5	6
3	1	4	2	5	3	6
4	1	5	2	6	3	4
5	1	6	2	3	4	5
6	1	2	3	5	4	6

Figure 2: An example of *index set*

4. Computational experiments

With the proposed π^{+} many calculations can be done in a parallel fashion on GPUs. For a meta-heuristic algorithm that handles the VRP on graphic processing units can easily calculate objective function and run local search operators.

4.1. Test Environment

The platform for our investigations is a modern NVIDIA Maxwell architecture GPU (GTX 980) with 16 streaming processors, each with 128 cores, a total of 2048 cores for parallel computation. In addition, our workstation has an Intel Xeon E5-2630 2.4 Ghz. CPU and 32 GB RAM. Python 2.7 was used when coding the algorithms. Furthermore, PyCuda [20] library is used for GPU operations. Finally, kernels are coded as C ++.

4.2. Performance of local search operators

As we do not design a total meta-heuristic algorithm for open VRP, we define a way to get performance of local search operators on GPUs by means of time. This should be considered that if we worked on an operator that has many conditional statement, the efficiency of GPUs would drop noticeably. Our performance evaluation way is easy. We are running intra-route (2-opt, Or-opt) and inter-route (one to one change) local search algorithms, respectively with different iterations. We do not allow vehicle capacities to be exceeded, therefore we can say that we have a time-consuming process during the control of vehicle capacities in the exchange between routes. Furthermore, we consider the time of initial solution transfer to GPU side. In contrast to objective function evaluation we handle a single solution.

Table 2: Performance comparison on local search operators

	A		B	
Iterations	**n80k10 CPU**	**n80k10 GPU**	**n386k47 CPU**	**n386k47 GPU**
1000	0.705	1.166	1.788	3.366
2500	1.655	1.754	3.670	4.785
5000	2.698	2.227	6.299	6.213
10000	5 993	3.566	14.365	9.137
25000	12.367	5.217	29.437	13.767
50000	22.335	7.425	56.432	18.998
100000	39.215	9.665	110.369	29.231

In Table 2, the results in seconds can be seen for different iterations for problem A and B. Up to 5000 iterations we can say that memory transfer is a bottleneck for GPU side. We can say that when the number of iterations increases, things turn in favor of GPU. As a consequence, a significant acceleration can be given to the algorithm with the proposed strategy.

5. Conclusions and further research

For high GPU utilization, there are a number of implementation issues, the most important ones being related to configuration and use of memory structures, and code diversion for configurable thread blocks. There is still a need to solve large combinatorial optimization problems like the VRP with fast meta-heuristic algorithms. GPU offers an opportunity to accelerate algorithms. In this study, we propose a novel permutation based solution representation technique which is easy to manipulate for the VRP. The proposed representation technique can be used with many meta-heuristic algorithms and many variants of the VRP. We showed the efficiency of this technique. For further studies, researchers can adapt the proposed representation technique to other combinatorial optimization problems such as quadratic assignment problem with new heuristic or meta-heuristic algorithms.

Acknowledgments: This work was funded with projects by Eskisehir Technical University, project numbers 1601F004 and 19ADP022.

6. References

[1] Gilbert Laporte, The vehicle routing problem: An overview of exact and approximate algorithms, *European Journal of Operational Research*, vol. 5 no. 3, pp. 345-358, 1992.

[2] G. Dantzig ve J. Ramser, The Truck Dispatching Problem, *Management Science,* vol 6, pp. 80-91, 1959.

[3] P. Toth ve D. Vigo, The granular tabu search and its application to the vehicle-routing problem, *Informs Journal on Computing,* vol. 15, no. 4, 2003.

[4] D. Pisinger ve S. Ropke, A general heuristic for vehicle routing problems, *Computers & Operations Research,* vol 34, pp. 2403-2435, 2007.

[5] S. N. Kumar ve R. Panneerselvam, A Survey on the Vehicle Routing Problem and Its Variants, *Intelligent Information Management,* vol. 4, pp. 66-74, 2012.

[6] E. Taillard, . P. Badeau, . M. Gendreau, . F. Guertin ve J.-Y. Potvin, A Tabu Search Heuristic for the Vehicle Routing Problem with Soft Time Windows, *Transportation Science,* vol 31, no. 2, pp. 101-195, 1997.

[7] K. Fleszar, I. Osman ve K. Hindi, Avariable neighbourhood search for the open vehicle routing problem, *European Journal of Operational Research,* vol. 195, pp. 803-809, 2009.

[8] Bin Yu, Zhong-Zhen Yang, Baozhen Yao, An improved ant colony optimization for vehicle routing problem, *European Journal of Operational Research*, vol. 196, no. 1, pp. 171-176, 2009.

[9] S. Tsutsui ve N. Fujimoto, Solving Quadratic Assignment Problems by Genetic Algorithms with GPU Computation: A Case Study, *GECCO*, Montreal, 2009.

[10] S. Tsutsui ve N. Fujimoto, Fast QAP Solving by ACO with 2-opt Local Search on a GPU, *IEEE Section Congres*, San Francisco, 2011.

[11] M. Czapinski, An effective Parallel Multistart Tabu Search for Quadratic Assignment Problem CUDA platform, *J. Parallel Distrib. Comput. ,* vol. 73, pp. 1461-1468, 2013.

[12] E. Özçetin, G. Öztürk, A Parallel Iterated Local Search Algorithm on GPUs for Quadratic Assignment Problem, *International Journal of Engineering Technologies*, vol. 4, no. 2, pp. 123-127, 2018.

[13] E. Özçetin, G. Öztürk, A Hybrid Genetic Algorithm for the Quadratic Assignment Problem on Graphics Processing Units, *Anadolu University Journal of Science and Technology-A Applied Sciences and Engineering*, vol. 17, no. 1, pp. 167-180, 2016.

[14] J. M. Cecilia, J. M. Garcia, A. Nisbet, M. Amos ve M. Ujaldon, Enhancing data parallelism for Ant Colony Optimization on GPUs, *J. Parallel Distrib. Comput. ,* vol. 73, pp. 42-51, 2013.

[15] A. Delevacq, P. Delisle, M. Gravel ve M. Krajecki, Parallel Ant Colony Optimization on Graphics Processing Units, *J. Parallel Distrib. Comput. ,* vol. 73, pp. 52-61, 2013.

[16] C. Shulz, Efficient local search on the GPU—Investigations on the vehicle routing problem, *J. Parallel Distrib. Comput.,* vol. 73, no.1, pp. 14-31, 2013.

[17] C. Groer, B. Golden ve E. Wasil, A Parallel Algorithm for the Vehicle Routing Problem, *INFORMS Journal on Computing,* vol. 23, pp. 315-330, 2011.

[18] J. Szymon ve Z. Dominik, Solving Multi-criteria Vehicle Routing Problem by Parallel Tabu Search on GPU, *Procedia Computer Science,* vol. 18, pp. 2529-2532, 2013.

[19] T. Davidović, P. Hansen, N. Mladenovic, Permutation-based genetic, tabu, and variable neighborhood search heuristics for multiprocessor scheduling with communication delays, *Asia Pacific Journal of Operational Research*, vol. 22, no.3, 2005.

[20] A. Klöckner, N. Pinto, Y. Lee, B. Catanzaro, P. Ivanov ve A. Fasih, PyCUDA and PyOpenCL: A scripting-based approach to GPU run-time code generation, *Parallel Computing,* vol. 38, no. 3, pp. 157-174, 2012.

Modelling of Financial Supports for SMEs and Evaluation of Alternative Policies with System Dynamics Approach

C.AKTAŞ[1]

[1] Agricultural and Rural Development Support Institution, Ankara/Turkey, caktas@hotmail com tr

Abstract **- In this study, the financial support payments for SMEs and developing alternative policies with the approach of "system dynamics" is discussed. The system dynamics approach is based on the examining complex systems as a whole and focusing on the dynamics of the system itself. For this aim, it is proposed that performing the impact analysis studies on how successful the public source used for supported firms to reach the determined targets and strategies not only through the outputs but also handling entirely the industrial systems and taking into account the complexity and feedback of the system through the system dynamics approach. In this context, the general structure of firms is modelled in terms of their internal dynamics, the interactions of these internal dynamics are revealed, and the effect of financial support on this system is observed**

Keywords **- System Dynamics, Modelling, Government supports, Impact analysis**

I. INTRODUCTION

Increasing the competitiveness of the global economy requires the efficient and effective use of resources as well as the highest value added created. Strategies for policy determination at regional level in the USA in the 1900s have revealed the concept of Development Agencies. Europe, on the other hand, has shaped this structuring peculiar to its own geography as a result of the emergence of economic and social differences between the regions since the second half of the 20th century. In Turkey, Development Agencies, as regional policy makers, have made a large transfer of public resources to SMEs since the year they started their activities by providing refundable and non-payment financing for the determined Projects and Activities. The implementation of the Development Agency model is about to complete its 10th year, and in this model, what kind of effect the non-refundable financing provided to SMEs has on businesses comes to the fore as an important research topic. In this regard, studies are carried out to measure the level of contribution and effects of financial support to companies, and the basic method in these studies is to compare the situation before and after the support.

II. SYSTEM DYNAMICS APROACH

The System Dynamics approach, which handles the complexity of the real life by including the interaction of the assets in the system with each other in the model, aims to give the most realistic result by considering the effect of the financial support provided by Development Agencies in the companies in general with a system approach. Especially in cases where the number of connections between entities is too high and cause-effect relations are continuity between system elements, in models where the relations between the parts of the system play a more dominant role than the individual effects of these parts, the elements that make up the system alone understand the behavior of the system as a whole. remains insufficient. The system dynamics method, on the other hand, allows complex systems to be taken from a holistic perspective, taking advantage of stocks, flows, feedback loops and time delays, and also allows observing the results of changes in the model, that is, evaluating the results of policy changes.

Jay Forrester first proposed the system dynamics approach in an article published in 1958 and the basic system logic was expressed. The aim of this article is to express and understand industrial processes with a sector approach. With the approach that aims to express the interactions and changes in the system in the best way by constructing the system with numerical data, cause and effect relationships are analyzed and the characteristic structure of the organization is revealed.

III. DEVELOPMENT AGENCIES AND GOVERNMENT SUPPORTS

The Development Agencies, for which the System Dynamics approach is the subject of this thesis; In order to ensure effective and efficient use of resources, to activate local potential, to accelerate regional development, to ensure sustainable development, to reduce inter-regional and intra-regional development differences, at least two and at most six provinces (Ankara, Istanbul and Izmir are exceptionally single provinces). It is the decision maker in a large economic and social system consisting of. Considering the complexity of the economic sectors as well as the social sectors within this system, System Dynamics can be put forward as a very powerful tool in order to determine the factors affecting the structure of the system as decision makers and to evaluate their interaction with each other. Within the scope of this study, the impact of Financial Supports, which is an important policy tool for economic sectors by Development Agencies, on companies will be discussed and evaluated with a system approach. Thus, decision makers will be able to handle all the complexity that the said Financial Supports will create on

companies and develop support mechanisms that will achieve the best results while allocating resources.

In Turkey, the private sector and in particular for SMEs, infrastructure, production, employment, R & D, financing, promotion and marketing, export and are given support in areas such as education. These supports are called intervention in general and the situation before and after the intervention is observed and the net effect is revealed. In determining the net effect, an evaluation will be made by considering the effects of displacement and tare loss.

Public financial supports will be handled with a system dynamics approach within the scope of this study and the effects of alternative policies on the real system will be evaluated. Therefore, these programs, which are implemented locally, are considered very important tools for accelerating regional development and thus contributing to national development.

1. Machine-Equipment Support
2. R&D Support
3. Human Resources Support
4. Quality Certificate Support
5. Promotion and Marketing Support
6. Institutional System / Software Support

IV. MATHEMATICAL MODELS

Stock variables are the order quantity waiting to be completed in the system, the products in stock are the current inventory quantity and the estimated demand quantity. In addition, variables such as the number of employees that do not appear directly in the model, accumulated money, main equipment, equipment inventory, sent products are also inventory variables. In the establishment system model, the levels that appear as the purchase orders being sent, the order amount in the correspondence process and the amount of the product in the transportation process to the company can also be taken as stock variables. The unit of measure for stock variables is the unit, except that the estimated demand is unit / week.

Flows of the created model; The order flow from the customer consists of the purchase orders sent, the flow of orders sent to the distribution sector, and the flow of products entering the inventory. The measure of flow variables is in units. Money flows that are not included in the model are also flow variables. The model generally includes information and product flows.

Covariates are expressed as subdivisions of decision variables. This is because they have independent meanings different from decision variables. The auxiliary variables are located in the information flow channels between levels and decision variables and are used to control the ratios in the system.

While each information line entering the auxiliary variables enables the determination of the variable to which the auxiliary variable is dependent; Each line that comes out refers to the transfer of information. Numerical values are not required for the auxiliary variable.

The auxiliary variables that appear in the model; The negative inventory control rate is the order rate intended to be shipped, the time taken to complete the orders, the amount of inventory intended to hold, the incomplete normal order quantity, the order quantity carried in the current flow, and the order quantity that must be supplied for operations. Rates; it is the week for the time taken to complete the orders, and for the others it is unit / week or unit.

Information flows are also of great importance for this institution in the retail sector. There are information flows between levels and covariates, from covariates to decision variables and therefore bilaterally between levels and decision variables.

Flow Variables in the Model:

Flow Variables; Stock variables are variables that show the changes that occur per unit time. Increases and decreases are transferred to the stock variables with the coefficient of variation. Since there is a proportional change in flow variables per unit time, flow variables in the model are shown with the expression "YD", which indicates the percentage change.

ProductionYD = Production TDO * Production Capacity (Change in Production Capacity in Unit Time)

Other Variables in the Model:

• Auxiliary: It is a dynamic variable calculated from other variables at a given moment. Utilities are often the most used type of variable. Expressions containing other variables can be included in the equation of an auxiliary variable.

• Constant: Its value does not change over time. Constants can be defined using the "GET XLS CONSTANTS" or "TABBED ARRAY" functions while taking the values on the right side of their equations.

• Data: It has variable values over time and is not dependent on model variables. If a variable is used in the model but is not defined, this variable will be treated as data variable.

• Group: Groups are not a real variable, they allow different variables to be used together.

• Beginner: A possible variable that other variables, such as the constant variable, can take on their first use.

• Lookup: It is the representation of nonlinear functions with numerical values. (Parameters x- and y-axis values)

• Character (String) Variable: Used as character string. MESSAGE function and Venapps are useful as tags.

• Units. Units are defined as additional information about the model variable and can be used to control the model by ensuring dimensional consistency.

QualityBS = (ProjeTB * KBDO) / QualityB Cost (auxiliary variable showing the Amount of Expenditure from the Project Budget for Quality Certificates)

QualityB Cost = 10000 (Constant variable showing the Average Expenditure for Each Quality Certificate)

ArgeDDO = (R&D Expenditures + ProjeTB * AHDO) / GrantButceTL (The variable that shows the increase of R&D expenditures to Project Support)

ExportDDO = 0.05 (Firm's Annual Export Rate Under Normal Situation)

ExportNDO = f (ProjeTB) (a function of the project budget depending on financial support)

Export TDO = ExportDDO + ExportNDO (Annual total change in exports)

Number of Staff (t) = Initial Value {t = 0} + ∫2_ $(t = 0)$ ^ $(t = n)$ $[\![(artim - azalıazal)\, dt,]\!]$

If expressed numerically;

Number of Personnel (t) = Personnel (t-dt) + (increment-decrement) * for dt, t = dt, 2dt ,,. Ndt

If this expression is generalized to any stock variable;

Stock (t) = Stock (t - dt) + (increment-decrement) * dt, t = dt, 2dt ,,. For ndt;

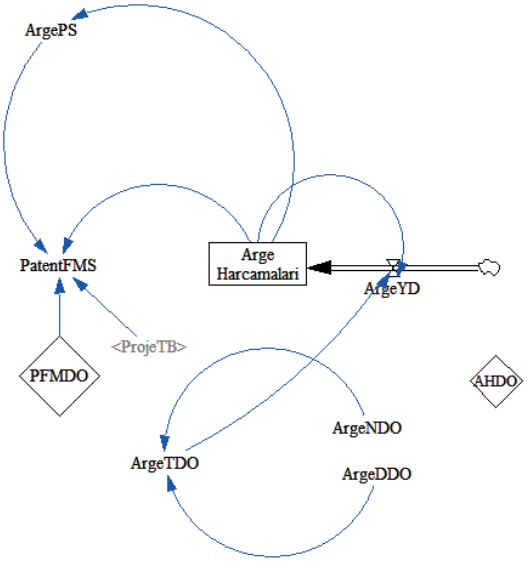


Figure:1 Model of R&D Expenditure In Systems

V. Conclusion

As a result of system dynamics approach for financial supports gives us dynamically observing the impact and additive effect created by public resources in real systems, evaluating the results of alternative policies in resource use, with the System Dynamics approach, obtaining effects and results at the system level, not only with the input-output dimension, evaluating the «multiplier effect» and «feedback» especially in social projects

References

[1] Jay W Forrester, World Dynamics Productivity Press, Portland, Oregon, 1989

[2] John Morecroft Strategic Modelling and Business Dynamics John Wiley &Sons Ltd, 2007

[3] Michael J Radzicki, Robert A Taylor A Systems Approach to Understanding Complex Policy Issues (1997)

[4] MIT Engineering Systems Division Working Paper Series ESD-WP-2003-01 13-ESD, Internal Symposium (2003)

[5] R G Coyle (auth)-System Dynamics Modelling_ A practical approach-Springer US (1996)

[6] Y Barlas Multiple tests for validation of system dynamics type of simulation models*, 1 European Journal of Operational Research, 42(1):59–87, 1989

Cloud Computing-based Time Series Analysis Using Google Colaboratory

Ö. YAKUT[1]

[1] Kocaeli University, Kocaeli/Turkey, onder.yakut@kocaeli.edu.tr

Abstract **- Nowadays, with the innovations and developments in technology, there is a great increase in the amount of data users use on the internet. For this reason, the need of various software and hardware tools to store, view, process and analyze this data on the internet is increasing day by day. Millions of data are produced around the world every day. To deal with this situation, Cloud Computing services are provided by leading companies around the world. In this study, time series analysis was performed using Google Cloud-based Google Colaboratory (Colab) service. In the Colab notebook, the Phyton programming language and the SARIMAX model, one of the machine learning methods was used. The well-known international airline passenger number data set was utilized as the time series. The time series was examined and the difference from the first order was taken and then transformed into a logarithm making it stationary (p-value <0.05). The time series (12 years) is divided into two for training (9 years) and testing (3 years). Then, the degrees of the terms belonging to the components of the SARIMAX model were determined and statistically evaluated and the most suitable SARIMAX(0,1,1)(0,1,1)12 model was selected. The SARIMAX(0,1,1)(0,1,1)12 model was trained with the training data set and then tested with the test data set. Performance criteria of the prediction results of the SARIMAX(0,1,1)(0,1,1)12 model are obtained as MAE 18.763, RMSE 22.876, MAPE 4.430% and R-squared 99.131%. The SARIMAX(0,1,1)(0,1,1)12 model developed according to these results has made highly accurate prediction. In this study, Phyton codes and results of time series analysis were recorded in Google Drive using a Colab notebook. While all these operations were done all we needed was a web browser and a computer running this browser. The hardware and software requirements of all analysis processes in the study were performed using Google Cloud Computing Services.**

Keywords **- Cloud Computing, Data Science, Google Colaboratory, Machine Learning, SARIMAX.**

I. INTRODUCTION

THESE days, with the increasing use of Internet-based technologies, the tendency to move the hardware and software needs to central servers and to run them over the web has increased. Users can obtain the services they need only through a web browser using the cloud computing platform they subscribe to. Thus, users save time and money for operations such as hardware and software investment, maintenance, configuration and updating.

The need for high-capacity hardware for analysis in areas such as Big Data, Machine Learning, Data Science and Computer Science, of which daily use is raising is increasing day by day. Cloud Computing makes the high-cost hardware requirements of the users more economical. The same situation also reduces software licensing costs as much as possible and enables the use of pirated software to be prevented. Cloud Computing provides great convenience for users to access the storage, code development, resource sharing, security, big data analysis and machine learning services they need at lower or no cost. Also, since Cloud Computing users store their data on the data centers of service providers, they can access this data whenever they need it from anywhere.

Another advantage of Cloud Computing technology is the convenience it provides to scientific researches. In this study, it enabled the web-based preparation and implementation of the machine learning algorithm developed for time series analysis. When this study is distributed on the web, researchers will be able to access the written code, analyze them according to the subject they will research, and record the developed machine learning method and analysis results. Then, researchers will be able to share the information and code obtained as a result of their operations with other authorized researchers. In the literature, there are some studies in various fields using Google Colaboratory.

Solano-Rojas et al. [1] run the densenet-121 architecture on the Google Colaboratory platform to detect Alzheimer's disease using Magnetic Resonance Images in their study. In the study, the limitations and consequences of using Google Colaboratory were mentioned. Hassan et al. [2] implemented the Convolutional Neural Network (CNN) method using the Google Colaboratory platform in order to process images and accurately predict the people in the images. Tan et al. [3] proposed a process line designed to analyze RNA-seq data on Google Colaboratory, which they called the Large-Scale Transcriptomic Analysis Pipeline in Cloud (LSTrAP-Cloud). Mark et al. In their study, [4] shared their experiences using Google Colaboratory, a hosted version of Jupyter Notebooks, in Artificial Intelligence (AI) courses at two universities. Cruz-Mendoza et al. in their study, [5] used a Long Short-Term Memory (LSTM) method using Google Colaboratory to estimate the number of people who will be infected in the following days with a dataset on the number of people infected with COVID-19 in Peru.

In this study, the time series analysis was performed using Google Cloud Computing based Google Colaboratory (Colab). The data set for the number of international airline passengers was used as a time series. In the analysis, the most suitable SARIMAX(0,1,1)(0,1,1)12 model from machine learning

methods was selected. When the performance result of this model is examined it is understood from the performance criteria that it has a high degree of accuracy. While performing these operations, the required hardware and software requirements were supplied from Google Cloud Computing services using Colab.

II. Material and Methods

A. Cloud Computing

As a result of the rapid increase in internet usage, cloud computing systems used in many areas are defined as an online storage system that provides functional convenience with web-based applications that do not require any installation. Some of the advantages of cloud computing technologies are as follows [6,7];

Flexibility and Efficiency: Cloud technologies allow access to the required hardware and software resources at the required level of need. Due to its flexible behaviour it saves individuals and companies from additional costs as it provides the opportunity to increase or decrease the capacity in case of need.

Availability: As long as there is internet network access, cloud computing systems make it possible to work without time and place dependency.

Security: In cloud computing systems, if individuals or organizations choose to store their information in cloud systems instead of their computers, it is possible that they can always access this information and continue their work more reliably.

Software Updates: Organizations need time and workforce to make various software updates. Since these workloads and costs are provided by cloud service providers in cloud computing systems, it provides companies with a serious cost advantage.

Eco-friendly (Less Carbon Footprint): It also makes a great contribution to preventing air pollution as energy consumption decreases due to the common use of resources on cloud computing systems by different companies.

Low Cost: Due to the common use of hardware and software resources in Cloud Computing platforms system costs are very low. The use of cloud infrastructure is very important as it can increase the profit margins of large subsidiaries and provide low-cost access to the resources needed for small companies and new entrepreneurs. As well as, there is no extra cost since the payment made is proportional to the resources used.

B. Cloud Computing Service Models

The service models used in cloud computing-based systems built by the National Institute of Standards and Technology (NIST) [7] are collected under three titles as in Fig. 1.

Software as a Service (SaaS) is a model that enables users to access internet service providers and cloud computing systems without any additional installation on their systems. In this model, the user's powers are limited and they cannot make changes on layers such as operating system, server, network. They only have management rights over the various settings in the services they use [7].

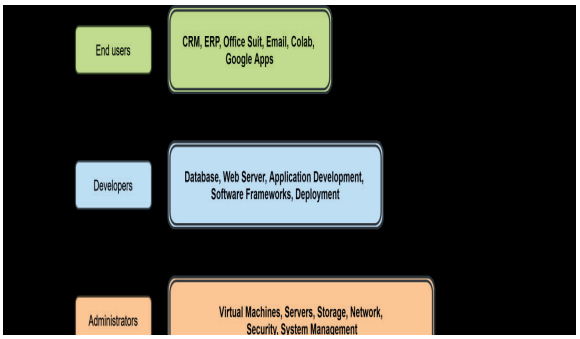


Figure 1: Cloud computing service models.

Platform as a Service (PaaS) is a service model that allows cloud computing system service providers to develop and run applications. This Platform provides the necessary support for the usage of the infrastructure and software services needed in the user's application. Here, the user has restricted their authority except for the components used only in own application [7].

Infrastructure as a Service (IaaS) is the model that the user can most frequently interfere with cloud computing services. In this model, the user has permission to make changes to hardware and software resources. In this model, the user can manage many structural changes such as installing the operating system according to the requirement, changing the registration areas, activating various local and web services and so on. As well as, the user can configure their security policies in this model if they wish [7].

C. Cloud Computing Deployment Models

According to the definition of cloud computing made by NIST, there are 4 different deployment models for service recipients [6].

Private cloud is cloud services provided for companies, holdings and similar users with services. Services and applications can be made by the company itself or by a different service provider. Private cloud computing offers companies control in the management of system resources.

Community cloud is the systems that have similar values such as security, policy and privacy of the cloud infrastructure and are shared with institutions or organizations acting together.

Public cloud is a cloud computing service provided over a public network. Public cloud service is available from a cloud service provider. The company performing the service provides software resources and makes it accessible to everyone on the internet. The system can be in its own data centers or different centers.

Hybrid cloud is a cloud environment with two or more different cloud distribution models. For example, a cloud consumer may prefer to use the public cloud for less sensitive cloud services while processing their sensitive data in cloud

services located within the private cloud.

D. Google Colaboratory (Colab)

Since the operations performed with Big Data in Machine Learning algorithms are too intensive, if your computer's processor speed, ram capacity and graphics card features are low, the analysis period can take days or even weeks. It is necessary to benefit from paid cloud services to make these analyzes in a shorter time. Google has provided a cloud computing service called Colab [8], which offers Graphics Processing Unit (GPU) and Tensor Processing Unit (TPU) support to its users free of charge. Colab allows users to improve their ability to write code. It also contributes to data analysis using Colab utilizing libraries such as TensorFlow, Phyton, PyTorch, Keras and OpenCV. In order to use the Google Colaboratory [8], you must create a new user account with Google or log in with an existing Google user account.

Firstly, a project folder is created in Google Drive for a new Colab notebook using a web browser. Afterwards, a new notebook is opened as in Fig. 2 by left-clicking the mouse in an empty area in this folder and selecting *Google Colaboratory*.

In the opened new notebook file user can change the name of the file. Also, under the *Runtime* menu, there are two options under *Change runtime type*. First, the user can choose the Phyton version (Phyton3) as the *Runtime type*. Secondly, the GPU and TPU can be selected as the *Hardware accelerator*. After this stage, the user can easily make their analysis using the GPU or TPU resource provided by Google. As a fundamental application given in Fig. 2, a sample Phyton code was written using the code cell in Colab notebook, NumPy to create data and Matplotlib to visualize data and a sine wave visual was created as output. You can transfer your data from Google Drive, GitHub and many other cloud-based sources to the Colab notebook. To analyze data, a machine learning model can be trained and you can evaluate and evaluate the model. A few lines of code that you type in the code cell in the Colab notebook are executed on Google's cloud servers and you get your data analysis results. Regardless of the power of your machine, users benefit from the power of Google hardware. In this case, all you need is a web browser to access the Google Cloud service.

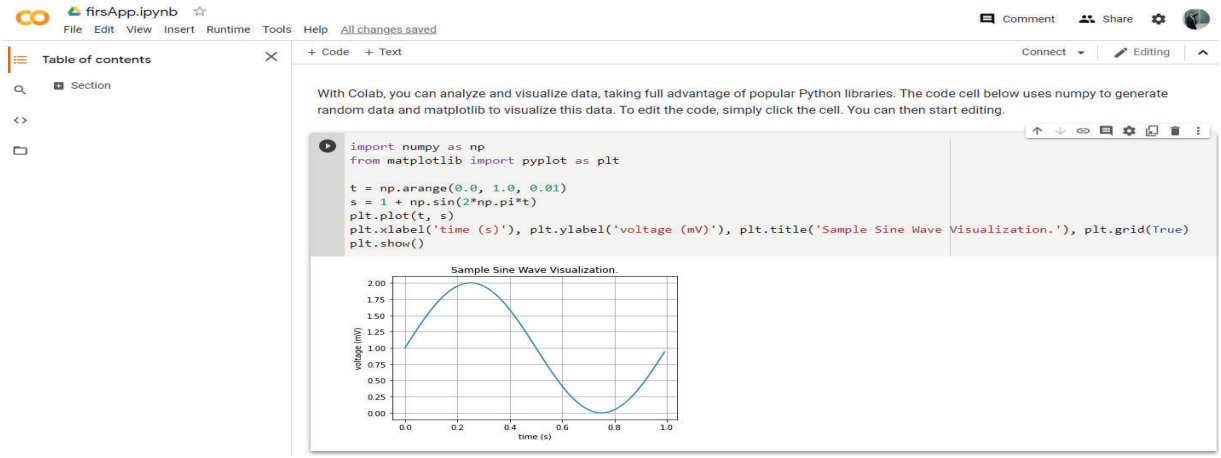


Figure 2: Colab notebook web browser view and a sample application.

E. Seasonal Autoregressive Integrated Moving Average with eXogenous (SARIMAX)

Seasonal Autoregressive Integrated Moving Average with eXogenous Regressors (SARIMAX) is a linear regression model used by including exogenous variables into the SARIMA (Seasonal Autoregressive Integrated Moving Average) model. The SARIMAX model is a model to be used in situations where the time series may show a seasonal trend or pattern.

External variables can also be added to the model via an external regressor term. The SARIMAX model with external regressors allows users to add the effects of external variables to the model. Exogenous variables are defined as variables that affect a model but are not affected by it [9].

The SARIMAX(p,d,q)(P,D,Q)s model consists of a combination of various terms. These terms those written with lowercase letters (p,d,q) represent non-seasonal degrees, those written with capital letters (P,D,Q) represent seasonal degrees and s represents the period in a season [10];

The explanation of these terms is given below;

- p is the order of the autoregressive (AR) term.
- d is the order of differencing operations
- q is the order of the moving average (MA) term.
- P is the order of the seasonal autoregressive (AR) term.
- D is the order of the seasonal differencing operations.
- Q is the order of the seasonal moving average (MA) term.
- S is the number of periods in a season.

F. Evaluation Metrics

In the equations given under this title, Y_t represents the real value, $\bar{Y}_t$ the predicted value, the relevant sample in the time t and n represents the total number of observations.

Mean Absolute Error (MAE) is adversely affected by data groups located at the endpoint. Therefore, the sum of the absolute values of the errors is averaged and a value is obtained. MAE is calculated with the help of Equation (1) [10].

$$MAE = \frac{1}{n}\sum_{t=1}^{n}|Y_t - \bar{Y}_t| \quad (1)$$

Mean Absolute Percentage Error (MAPE) is found with the help of Equation (2). This approach is useful in assessing the accuracy of the forecast. MAPE shows how large the predictive errors are when compared with actual values. [10].

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\frac{|Y_t - \bar{Y}_t|}{Y_t} \times 100 \quad (2)$$

Root Mean Square Error (RMSE) is a quadratic metric that is often used to find the distance between the predicted values and actual values of a machine learning model. RMSE is the standard deviation of the estimation errors (residues).

In other words, it tells the user how dense that data is around the line that best fits the data. It is calculated with the help of Equation (3) [10].

$$RMSE = \sqrt{\frac{\sum_{t=1}^{n}(Y_t - \bar{Y}_t)^2}{n}} \quad (3)$$

R-squared is the criterion that represents predictive performance for linear regression models. R-squared explains the percentage volume of cases where the independent variable affects the dependent variable. It shows the strength of the relationship between the independent variable and the dependent variable in the range of 0-100%. As the R-squared value approaches 100% the user can say that the relationship is very strong. In this case, the accuracy rate of the predictions to be made through the model will also perform well.

III. Experimental Study

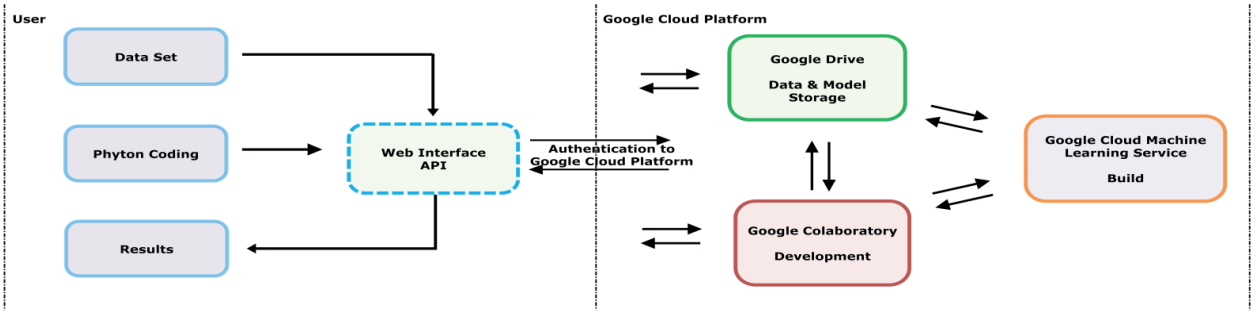


Figure 3: Google cloud computing-based time series analysis using Google Colaboratory.

In this study, a method is proposed to perform time series analysis using the airline passenger count data set as in the block diagram in Fig. 3. The user accesses to the Google account by sign-in Google via the web browser. In this way, he can benefit from all the services provided by the Google Cloud Platform. The user accesses the Google Colab web page via the internet address [8]. A Colab notebook for the user is opened here. It changes the name of the opened notebook and saves it in the Google Drive folder. The user can also upload the airline passenger count data set in the project folder in Google Drive to the storage area of the session using the Files section in the Colab notebook. Thus, the data set becomes accessible by the Colab notebook. The developed Phyton code is written into the code cell in the Colab notebook and then the Run Cell command is given. After that, Google Colabratory transmits the written Phyton code to the Google Cloud Machine Learning Service. Phyton code is compiled and built here. If an error occurs as a result of this process, this error is displayed to the user via the Colab notebook. If there is no error, the result is displayed in the output section of the Colab notebook and the analysis process is also saved in the project folder in Google Drive.

If the analysis result is positive, the machine learning model developed for time series analysis can be deployed over cloud-based storage systems such as Google Drive or GitHub.

IV. Experimental Results

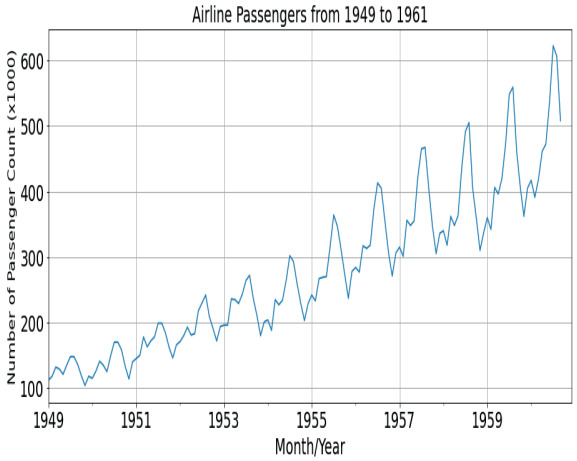


Figure 4: Airline passengers from 1949 to 1961.

In this study, the data set used as a time series includes the monthly count of international airline passengers from 1949 to the end of 1960. This data set has been added to the Colab notebook working environment in the project folder in Google Drive. Distribution of the passenger's counts in the data set over time is shown in Fig. 4.

Before Transformation		After Transformation	
Dickey-Fuller Test:		Dickey-Fuller Test:	
Test Statistic	0.815369	Test Statistic	-3.162908
p-value	0.991880	p-value	0.022235
Lags Used	13.000000	Lags Used	13.000000
No. of Obs	130.000000	No. of Obs	119.000000
Critical Value (1%)	-3.481682	Critical Value (1%)	-3.486535
Critical Value (5%)	-2.884042	Critical Value (5%)	-2.886151
Critical Value (10%)	-2.578770	Critical Value (10%)	-2.579896
(A)		(B)	

Figure 5: Augmented Dickey-Fuller test results; (A) before the transformation, (B) after the transformation.

Time series variables should be tested for stationarity. In the study, the Augmented Dickey-Fuller (ADF) test was applied to the data set. The result of the test statistic is shown in Fig. 5. It is understood that the data set is not stationary since the p-value is higher than 0.05 (p-value=0.991) as seen in Fig. 5(A).

In order to stationarize the time series, first-order difference and transform to logarithms were made. As a result of this process, the Augmented Dickey-Fuller (ADF) test was applied again. It is understood that the data set became stationary because the p-value was less than <0.05 as seen in test statistic shown in Fig. 5(B).

It is necessary to select the most suitable model by determining the degrees of the terms of the components of SARIMAX which is the machine learning method to be developed for time series analysis. At this stage, the model with the lowest Akaike Information Criteria (AIC) value was selected among the developed models. Statistical results of the selected SARIMAX(0,1,1)(0,1,1)$_{12}$ model is shown in Fig. 6.

```
*** Finding the model with minimum AIC value ***
param              (0, 1, 1)
param_seasonal  (0, 1, 1, 12)
AIC                   895.681
Name: 27, dtype: object
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
ma.L1         -0.3237      0.072     -4.485      0.000      -0.465      -0.182
ma.S.L12      -0.1044      0.096     -1.089      0.027      -0.292       0.083
sigma2       143.4990     17.896      8.019      0.000     108.424     178.574
==============================================================================

==============================================================================
Ljung-Box (Q):                       53.39   Jarque-Bera (JB):                 4.04
Prob(Q):                              0.08   Prob(JB):                         0.13
Heteroskedasticity (H):               2.32   Skew:                             0.14
Prob(H) (two-sided):                  0.01   Kurtosis:                         3.88
==============================================================================
```

Figure 6: Finding SARIMAX model with minimum AIC value.

Data Set contains 12 years of data. Data set is divided into two as training (9 years) and testing (3 years). The training result of the best fit SARIMAX(0,1,1)(0,1,1)$_{12}$ model determined for time series analysis is shown in Fig. 7. The statistical results of the model are presented in Table 1.

Table 1: Prediction results of the SARIMAX(0,1,1)(0,1,1)$_{12}$ model

	Value
MAE	18.763
RMSE	22.876
MAPE (%)	4.430
R-Squared (%)	99.131

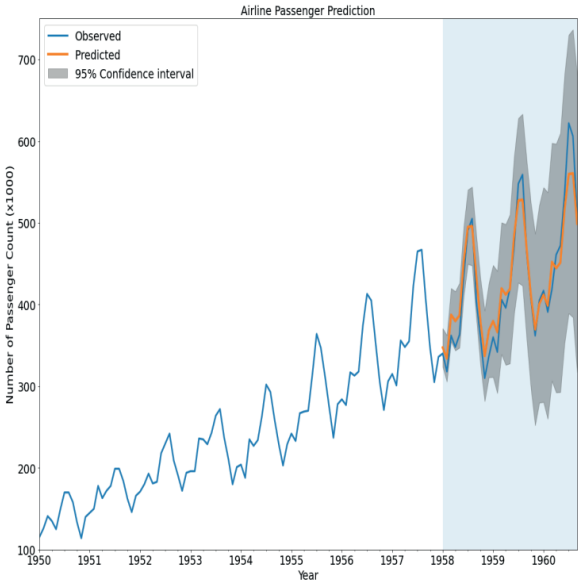


Figure 7: Prediction result of the SARIMAX model obtained using the training data set.

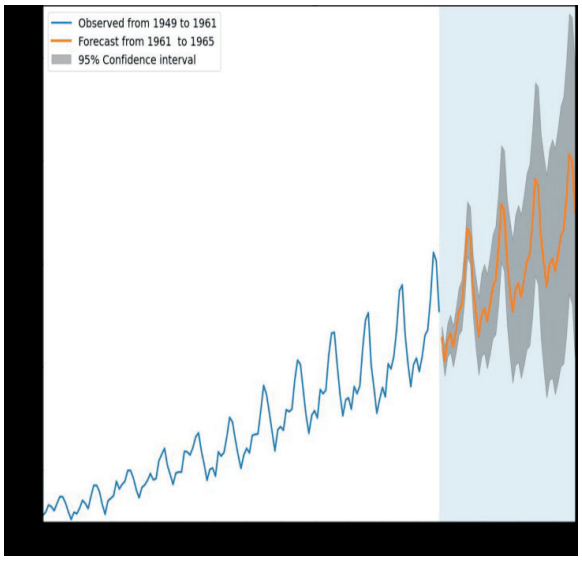


Figure 8: Forecast result of the SARIMAX(0,1,1)(0,1,1)$_{12}$ model.

For time series analysis, the data set was stationarized and the SARIMAX$(0,1,1)(0,1,1)_{12}$ model was created, trained, and used to make a forecast. The SARIMAX$(0,1,1)(0,1,1)_{12}$ model forecasted passengers count for the next 4 years (48 months) with a 95% confidence interval. The forecast result of the model is shown in Fig. 8.

V. Discussion

In this study, the count of airline passengers is forecasted with Google Cloud Computing based Colab notebook. Therefore, time series analysis was performed using the SARIMAX method developed with the Phyton programming language. The SARIMAX model with the lowest AIC value was chosen. The degree of the components belonging to the selected SARIMAX model was determined as SARIMAX$(0,1,1)(0,1,1)_{12}$.

Selected SARIMAX$(0,1,1)(0,1,1)_{12}$ model was trained using time series data reserved for training. Next, the model is tested with the time series data reserved for testing. As a result of this process, the low MAPE value of 4.430% (MAPE <10%) indicates that the success of the model in predictions for the future is significantly high.

The correlation coefficient of the R-squared value is 99.131%. The R-square value indicates that the model is determined as the best fit model. The MAE value is 18.763 which shows the proximity of the prediction value to the observed value. The MAE shows that the average size of the model's prediction error is low. The RMSE value of the model is 22.876. Since the difference between the RMSE and the MAE value is small (4.113), the size of the error involved in the prediction made by the model is less.

By using the SARIMAX$(0,1,1)(0,1,1)_{12}$ model proposed in this study, the forecast was made for the count of airline passengers for four years. As a result of the forecast made, it is concluded that the count of airline passengers will continue to increase each year. This result can be used to help policymakers make policy adjustments and invest in order to better serve passengers and the industry, to recognize potential obstacles, and to take new opportunities and lead.

Colab notebook offers a cloud-based service that facilitates user work such as learning software, working in data science, researching artificial intelligence or machine learning. While using this service, the user only needs a web browser and an electronic device (such as a computer, tablet, smartphone) running this browser. The developed methods or algorithms run using the Google Cloud infrastructure. So, users save time and money on software and hardware requirement and management.

VI. Conclusion

In this study, Google Colaboratory or Colab notebook which is based on Google Cloud Computing was used. The SARIMAX machine learning method developed in the Phyton programming language was utilized in the Colab notebook. International airline passenger count data set was used for the time series analysis with SARIMAX machine learning method. This data set has been analyzed in terms of stationarity and the data set has been stationarized using various tests and methods. Afterwards, the SARIMAX model was selected and as a result of the SARIMAX$(0,1,1)(0,1,1)_{12}$ model which has the lowest AIC value was determined. This model has been trained and tested with the data set. Performance results of the model are MAE 18.763, RMSE 22.876, MAPE 4.43% and R-Squared 99.131%. In the evaluation, it was concluded that the prediction ability of the developed model was highly accurate. Then, a four-year forecast was made using this model. In this study, time series analysis has been made using a cloud computing-based machine learning model. The software and hardware required for this process were accessed via a web browser and this open-source ecosystem was used. In future, Colab notebook will be used effectively in areas such as image processing, signal processing and natural language processing with machine learning algorithms.

References

[1] Solano-Rojas, B , Villalón-Fonseca, R , & Marín-Raventós, G , "Alzheimer's Disease Early Detection Using a Low Cost Three-Dimensional Densenet-121 Architecture", *In International Conference on Smart Homes and Health Telematics 2020*, pp 3-15

[2] Hassan, N I , Tahir, N M , Zaman, F H K , & Hashim, H , "People Detection System Using YOLOv3 Algorithm", *In 10th IEEE International Conference on Control System, Computing and Engineering 2020*, pp 131-136

[3] Tan, Q W , Goh, W , & Mutwil, M , "LSTrAP-Cloud: A User-Friendly Cloud Computing Pipeline to Infer Coexpression Networks", *Genes*, vol 11(4), pp 428, April 2020

[4] Mark J Nelson and Amy K Hoover , "Notes on Using Google Colaboratory in AI Education", *In Proceedings of the CM Conference on Innovation and Technology in Computer Science Education 2020*, pp 533-534

[5] Cruz-Mendoza, I , Quevedo-Pulido, J , & Adanaque-Infante, L , "LSTM perfomance analysis for predictive models based on Covid-19 dataset", *IEEE XXVII International Conference on Electronics, Electrical Engineering and Computing 2020*, pp 1-4

[6] Peter Mell, Timothy Grance, "The NIST Definition of Cloud Computing", *NIST Special Publication 800-145*, pp 1-7, September 2011

[7] Cloud Computing, Available: https://en wikipedia org/wiki/Cloud_computing#:~:text=Proponents%20also%20claim%20that%20cloud,the%20burst%20computing%20capability%3A%20high, [Accessed 09 10 2020]

[8] Google Colaboratory, Available: https://colab research google com/, [Accessed 10 10 2020]

[9] Cools, M , Moons, E , & Wets, G , "Investigating the variability in daily traffic counts through use of ARIMAX and SARIMAX models: assessing the effect of holidays on two site locations", *Transportation Research Record*, vol 2136(1), pp 57-66, January 2009

[10] C Chatfield, Time-series Forecasting Chapman & Hall CRC, 2001

Analysis of Cryptocurrency Price Changes with Big data Tools

Z.KAMIŞLI ÖZTÜRK[1] and F. BAYRAKTAR[1]

[1] Eskisehir Technical University, Eskişehir/Turkey, zkamisli@eskisehir.edu.tr
[1] Eskisehir Technical University, Eskişehir/Turkey, furganbayraktar@gmail.com

Abstract **- Basic and technical analysis methods are used in the determination of trading strategies in stock exchanges. While determining the trading strategies with technical analysis, very large data are encountered. This situation makes it difficult to analyze the data and increases the likelihood of incorrect strategies. Cryptocurrencies is one of the commodities used for trading transactions, and people who are unable to analyze these big data correctly may face huge losses from cryptocurrency investments. Due to the recent emergence of cryptocurrencies, there is no method and approach in the literature to analyze changes in cryptocurrency prices. In this study, data of the 3 cryptocurrencies with the largest market value are used. These cryptocurrencies are Bitcoin, Ethereum and XRP. For the analysis of price changes of these cryptocurrencies, K Nearest Neighbors (KNN), Support Vector Regression (SVR) and Random Forest (RF) algorithms were used. As a result of the study, the estimated data is compared with the actual data. It was seen that the Support Vector Regression gave the most successful result in all data sets.**

Keywords **- Big data, Cryptocurrency, K Nearest Neighbors, Random Forest, Support Vector Regression.**

I. INTRODUCTION

TODAY'S technological developments causes different investment instruments. The blockchain technology that emerged as a result of these developments forms the infrastructure of cryptocurrencies. Crypto money is digital and virtual money that is encrypted using the science of cryptology and can be used in payment and money transfer transactions. Bitcoin (BTC) is the most known of the crypto currencies, was created in 2009. Bitcoin features a point-to-point distributed network. All transactions taking place on this network are recorded and these records collected are included in structures called blocks. A specific expression must be found by applying hashing algorithms that require very large operations on each of these blocks. The first user to make this statement is rewarded with a specific BTC. The prize amount is reduced by half for every 210,000 blocks. This way of generating BTC by this method is called mining. With this method, BTC can be produced by using the processor power of computer hardware, without being connected to a center, unlike normal currencies. It has started to be preferred by people because it is not connected to a certain center, its supply is fixed, and it is cheap, fast and difficult to follow for money transfer between users. As BTC became widespread among people, it was used as a payment tool, and it was also seen as an investment tool with the thought that its price would increase in the future.

There are many crypto currencies besides BTC. One of these crypto currencies is Ethereum (ETH). ETH is in the second place in the crypto money market in terms of market value. ETH is a platform that uses blockchain infrastructure like Bitcoin and allows users to create decentralized applications. Another crypto currency to be analyzed in this study is XRP. Unlike Bitcoin and Ethereum, XRP is a centralized cryptocurrency and its maximum supply in the market is 100 billion units. Like XRP, Bitcoin and Ethereum, they are not mined.

As in all exchanges, basic and technical analysis methods are used when determining the trading strategies in crypto money exchanges. Fundamental analysis is a type of analysis that is used to predict future price movements of a commodity, based on information and news related to supply and demand statistics, economic, political, environmental and other factors. Technical analysis can be explained as trying to predict future price changes by examining past price movements. While conducting technical analysis, very large data are encountered. In this case, the investor has to determine the trading strategies without knowing which of these data are meaningful or which are meaningless. As a result, the strategies to be implemented are more likely to be erroneous and the investor may suffer from this.

Alkhatib et.al. [1] estimated the stock prices traded on the Lebanese stock exchange using KNN and non-linear regression methods. İlarslan [2] made the price prediction of stocks in Borsa İstanbul (BİST) using the KNN algorithm. Filiz et.al. [3] aimed to classify the factors affecting the direction of the BIST-50 index on a daily time basis by using KNN, Naive Bayes, C4.5 algorithm and Artificial Neural Networks. In Lahmiri and Bekiros [4], prediction of high volatility price series such as Bitcoin with machine learning methods was compared. The data set was analyzed using SVR, Gaussian Poisson Regression, Regression Trees, KNN, Bayes Regulatory Neural Network, Feed Forward Neural Network and Radial Based Function Neural Networks. Yalçın (2016) estimated the instantaneous fuel consumption of aircraft during the flight using SVR, decision trees, Artificial Neural Networks and Gaussian Process methods. The obtained data was modeled as a nonlinear function and this model was tested on a real data set. In Fradinata et.al. [5], another study using the SVR method, it was aimed to estimate the noodle demand. The SVR method was also used in Meral [6] to estimate the monthly natural gas consumption amount of power plants. In the study of Ince and Trafalis [7], it was tried to predict the price movements of

stocks with Multi-Layer Detector, SVR and Integrated Autoregressive Moving Average methods.

Herrera et.al. [8] used a hybrid model using traditional economic methods, Artificial Neural Networks and RF methods in their study to make long-term price estimates using monthly prices of energy products. In another study, Baba and Sevil [9] predicted the returns of the public offering by RF algorithm.

Kumar et.al. [10] used SVM, RF, KNN, Naive Bayes, and Softmax algorithms for price estimation of stocks. Mallqui and Fernandes [11] made estimates of the daily highest price, daily lowest price and daily closing price with Artificial Neural Networks and SVM methods by using daily data of Bitcoin in US Dollars (USD) basis. RF, SVR, and Logistic regression algorithms were also used in Chen et.al. [12] in order to predict Bitcoin prices.

When many studies in the literature are examined, it has seen that K Nearest Neighbor (KNN), Support Vector Regression (SVR) and Random Forest (RF) algorithms are used in price estimation and other forecasting applications in finance, industry and service sectors. In these studies, in the field of finance, it has been observed that price estimations are generally made using BIST and stocks in various foreign exchanges. Although there are different studies in the studies on crypto money in the literature, it has been seen that there are generally only studies on BTC.

In this study, different from the literature, data sets of three different cryptocurrencies named BTC, ETH and XRP were created by taking into account the daily data of the crypto currencies from certain stock exchanges to the present, and these data sets were aimed to make price estimation using all KNN, SVR and RF algorithms. It is very important for the contribution of this study to the literature that the crypto money price estimation using these three methods together and then finding the most successful method for the real price value is very important.

II. Methods Used

In this study, it is aimed to estimate the price of Bitcoin (BTC), Ethereum (ETH) and XRP, which are the three most known cryptocurrencies with the largest market capitalization, on USD basis.

A. Data Set

All three data sets to be used in this study were taken from https://coinmarketcap.com/ on the basis of daily price changes. Table 1 shows the dates on which each data set consists of crypto data. The data sets consist of six continuous valued features: daily opening price, daily closing price, daily highest price value, daily lowest price value, daily transaction volume, and daily total market value of crypto currency.

The continuous increase in the number of data in the data set used in this study, the fact that the data obtained from the Internet environment is not structured, the data size increases depending on the data produced day by day, the obtained data set is obtained from coinmarketcap.com, a reliable source about cryptocurrencies, and considering the financial gain to be obtained as a result of the analysis of these data, this data set covers five factors of big data.

Table 1: This caption is centered.

Data Set	Opening date	Closing data
BTC	27.12.2013	23.04.2020
ETH	07.08.2015	23.04.2020
XRP	01.01.2014	23.04.2020

Since there is no daily transaction volume data in the BTC and XRP data set, the cleaning process was carried out while preparing this data set.

After the BTC, ETH and XRP data sets are cleared, min-max normalization is performed using Equation (1), where x_i is the attribute value in the i^{th} row.

$$\frac{x_i - \min(x_i)}{\max(x_i) - \min(x_i)} \quad (1)$$

After this normalization process, KNN, DVR and RF methods were applied to all data sets using WEKA 3.8 [13] program.

The data to be used were tested with *k*-fold cross validation method. In the *k*-fold cross validation method, firstly a *k* value is determined. The data set is divided into *k* sub-data sets of equal size. One of these sub-data sets is called test data set and the remaining *k*-1 sub-data set is defined as the training data set. Training data sets are data stacks consisting of data presented for the learning of the algorithm used. Test data, on the other hand, are data stacks that are used according to the model created after the algorithm has finished learning, and to test how close the model gives predictions to the real data. In K-fold cross validation, the model is trained using one test data set and *k*-1 training data set after the whole data set is divided into k pieces. During *k* number of iterations, the model is trained by changing the test data in each iteration.

As a result of the algorithm, the model giving the lowest mean absolute error (MAE) of the test data sets is used as the ideal *k* value. MAE value is calculated using Equation (2) where e_j is the difference the difference between the predicted and real values. It can be said that the closer the MAE value is to zero, the better the predictive performance of the model.

$$MAE = \frac{1}{k} \sum_{j=1}^{k} |e_j| \quad (2)$$

B. K Nearest Neighbor (KNN)

K Nearest Neighbor Algorithm (KNN) is one of the methods used in classification and regression studies in supervised learning algorithms. Unlike other supervised learning algorithms, it does not have a training phase. The training phase and the testing phase are almost the same. Since it is a lazy learning type, it is an algorithm that is not ideal to implement in large data sets. The class of the data added to the KNN sample data set is determined by calculating the distances to the existing data in the data set and looking at the specified number of neighbors. Distance calculations for continuous variables are

usually calculated using one of the Euclidean, Manhattan and Minkowski distance functions.

The steps of the KNN algorithm are briefly as follows:

Step 1: As a result of the sample data set analysis, the k value that gives the least error is found.

Step 2: The distance of the new data to be added to the existing data set is calculated by using the distance function.

Step 3: After calculating the distances, the nearest k number of neighbor data to this data are considered.

Step 4: Assigned to the nearest k neighbor class according to attribute values.

C. Support Vector Regression (SVR)

Support Vector Machines algorithm is one of the best kernel-based methods used to solve classification and regression problems. This very powerful learning strategy based on machine learning algorithm principles was developed by Vapnik [14].

The expression given as $\{(x_i, y_i)\}_{i=1}^{n} \subset R$ of the training data also shows the area of the X input data. This X variable can be daily price data in crypto coins. Y variable is defined as the prediction variable in regression problems. The aim of the SVR is to find the $f(x)$ function with the highest deviation from the targets obtained for all training data and at the same time as flat as possible. As can be understood from this expression, as long as errors are less than ε, they will not be ignored by the method, but any deviation greater than ε will not be accepted.

D. Random Forest (RF)

Random Forest (RO) algorithm is a bagging-based machine learning method. The RO algorithm creates multiple decision trees, combining them to create much more accurate and stable models. The RO algorithm is also used for regression analysis. When used in regression analysis, the target class consists of continuous data. Therefore, each tree created by the RO algorithm creates a continuous prediction model for us. The estimate of the whole model is obtained by taking the average of the estimates obtained from individual trees [15]. The steps of the RO algorithm can be listed as follows:

1. N pieces of data are selected from the data set and a certain part of these selected data is used as learning data and trees are formed.
2. m number of attributes are selected from the learning data set and the one that will gain the most knowledge is determined. The ideal value for the regression tree is $m = {}^{1}/_{p}$, and the ideal value for classification is $m = \sqrt{p}$, where p indicates the total number of features.
3. It is calculated by using the best branching criterion of the chosen prediction variable, entropy and Gini index. According to this value, each node of the data set is divided into two sub-branches.
4. The processes up to this point are repeated until a leaf knot is obtained.
5. The previous steps are repeated N times to create N number of trees.
6. The model is tested with test data using the rules created by these trees.
7. For classification problems, the tree with the most votes is chosen as the estimate of the entire forest. For regression problems, the average of all voting is calculated and estimated accordingly.

III. Estimation of Cryptocurrency Price Changes With KNN, SVR and RF

In the study, price estimation of defined cryptocurrencies (BTC, ETH, XRP) was made using the WEKA 3.8 with KNN, SVR, and RF methods for three different data sets created using six different features.

As given in Section II, in order to determine the model to be created for the KNN, SVR, and RF methods, the BTC, ETH, and XRP training data sets were created by changing the k-fold cross validation values, and the model with the lowest MAE value was selected as the prediction model for the related cryptocurrency.

A. BTC Price Estimation

In order to apply the KNN algorithm to the Bitcoin data set, firstly, using the separated training data set, the lowest MAE value and the k-fold cross validation value were found. Due to the space limit, the k-fold crossover results are given only in the BTC price estimation using the KNN method in Table 2.

Table 2: k-value and k-fold cross validation results of Bitcoin KNN models.

	***k* times cross validation MAE value**					
k	**5**	**6**	**7**	**8**	**9**	**10**
1	167,9031	169,6786	167,1244	167,2375	167,9599	167,4398
3	137,4052	139,9591	136,6819	138,0851	137,8709	137,6872
5	131,9727	132,8269	130,6864	131,5973	131,6963	131,3264
7	127,4385	129,8612	128,0758	128,7563	128,7736	127,7193
9	125,9282	127,1355	125,579	124,8323	126,7831	126,7391
11	126,1409	126,6652	125,1089	**124,8286**	125,0342	124,8712

For Bitcoin, it is seen that the lowest MAE result is obtained when the k parameter is 11 and the k-fold cross validation value is 8. Price estimation has been made using these parameters. The price estimation, error value and MAE result made with KNN are given in Table 3.

Table 3: BTC price prediction analysis results made with the KNN. *method*.

Real value	**Estimated Value**	**Error Value**
6793,62	6479,698	-313,922
6733,39	6749,171	15,781
7271,78	7061,659	-210,121
7176,41	6615,73	-560,68
⋮	⋮	⋮
7550,9	6887,646	-663,254
MAE		260,0715

When the results of the KNN method were examined, it was seen that the prediction success obtained at some points did not spread to the general. When the mean absolute error value of the estimated values is examined, an error rate of nearly 4% was observed.

When the training set was analyzed with the SVR method, the lowest MAE value was obtained when the *C* coefficient was selected as 16 and the *k*-fold cross-validation value was selected as 5. Price estimation has been made with the model obtained using these values. It has been observed that the error rate in BTC price prediction with the SVR method has dropped to 2% -2.3%. When analyzed on the basis of MAE value, more successful estimation results were obtained than KNN method.

In order to determine the model to be created for the RF method, the BTC training data set was created by changing the *k*-fold cross validation values and the models with the lowest MAE value were selected as the prediction model. When the estimated values obtained by RF algorithm were examined, it was seen that the results obtained were more successful than the results obtained in the KNN method. However, compared to the SVR method, it produced more erroneous predictive data.

As a result of analyzing the BTC data set with KNN, DVR and RF methods, it was concluded that the most successful method in this data set is the SVR method. In Figures 1, 2 and 3, the estimated data obtained in the methods are compared with the real data.

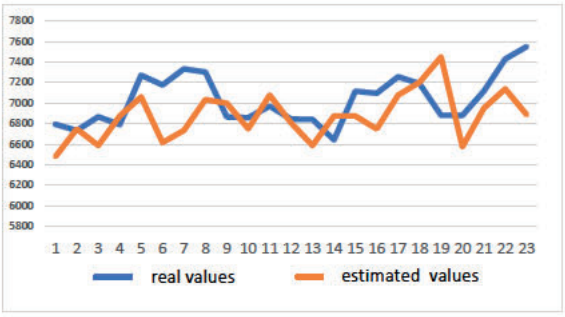


Figure 1: Comparison of BTC data obtained by the KNN method.

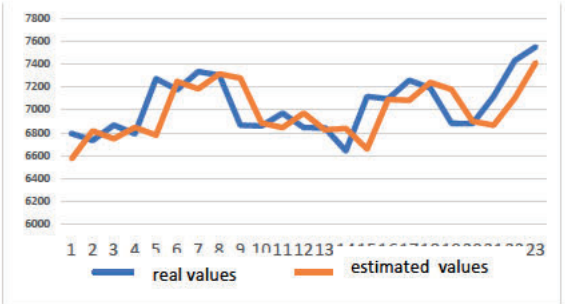


Figure 2: Comparison of BTC data obtained by the KNN method.

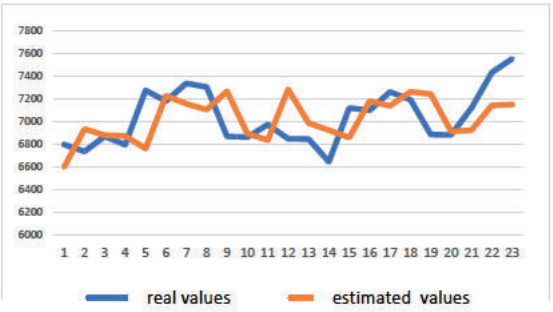


Figure 3: Comparison of BTC data obtained by the RF method.

It is seen that the values of the SVR algorithm, which gives the most successful results, act in parallel with the real values, reacting faster to price changes compared to other methods. In the RO algorithm, there were serious deviations at some points and its reactions to price changes were slower. In the KNN graph, it can be said that it failed both in terms of the data it generates and in terms of its response to change, according to the other two algorithms.

B. ETH Price Estimation

As a result of the analysis of the ETH training data set with the KNN method, it was seen that the lowest MAE result was obtained when the *k* parameter was selected as 9 and the *k*-fold cross validation parameter as 10. When the estimated values obtained using the KNN method of ETH price estimation are examined, it was observed that the deviations from the real value have very high values. When the error values of the estimated values obtained by the model were examined, an average error rate of more than 8% was observed.

It was obtained that the average error rate of Ethereum price estimation with the Support Vector Regression algorithm is 3%. Unlike the BTC price prediction of the same method, the MAE value obtained from the ETH price prediction scored lower than the MAE value of the training set. From this perspective, the SVR method has been satisfactorily successful for ETH price prediction.

When the estimation values obtained by the Random Forest algorithm were examined, it was calculated that the average error rate is around 4%. Its success is lower than the SVR method. It gave more successful results than the KNN algorithm. As in the SVR method, a lower mean absolute error value was obtained from the MAE value obtained from the training data in the RF algorithm.

Based on all these results, it can be said that the most successful method in Ethereum price prediction analysis is the SVR algorithm. Comparisons of data obtained by KNN, SVR and RF algorithms with real ETH data are given in the graphics in Figures 4, 5 and 6.

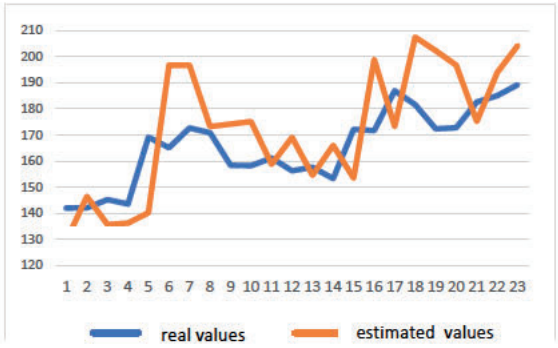


Figure 4: Comparison of ETH data obtained by the KNN method.

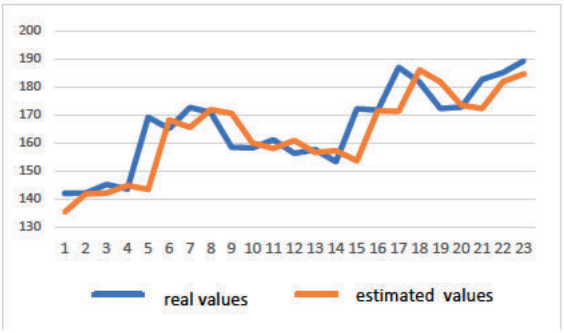


Figure 5: Comparison of ETH data obtained by the SVR method.

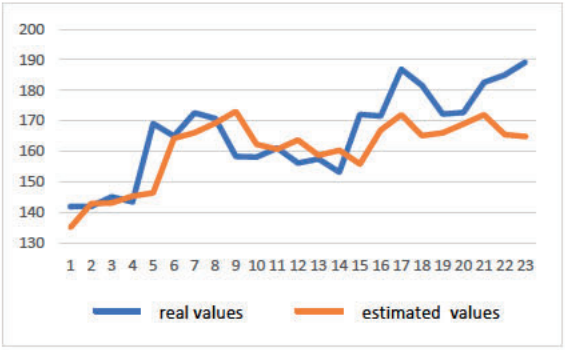


Figure 6: Comparison of ETH data obtained by the RF method.

C. *XRP Price Estimation*

In order to use the KNN algorithm, the XRP data set is first divided into training and test sets. In the XRP data set, the parameter pair with the lowest mean absolute error score was obtained when the k parameter was selected as 11 and the k-fold cross validation parameter as 10. When the estimated values of XRP price estimation obtained by the KNN algorithm are examined, it is seen that the average deviations from the real values are realized around 5%. Considering the average error rates of three data sets with the KNN method, this method gave the most successful results in the XRP data set.

It is seen that the average error rate of XRP price estimation using the DVR algorithm is 2.2%. The average absolute error value of the data estimated by the SVR method scored 60% less than the average absolute error value of the training set. Thus, this method showed the most successful price forecast in the XRP data set.

When the XRP prediction data were examined, it was calculated that the average error rate was 2.3%. When the SVR and RF algorithms are compared on the basis of the XRP data set, it is seen that they give very similar results. The KNN algorithm gave less successful results compared to these methods. As in the SVR method, a lower mean absolute error value was obtained from the MAE value obtained from the training data in the Random Forest algorithm. In the graphics in Figures 7, 8 and 9, the comparisons of the data obtained by KNN, SVR and RF algorithms with real data are given.

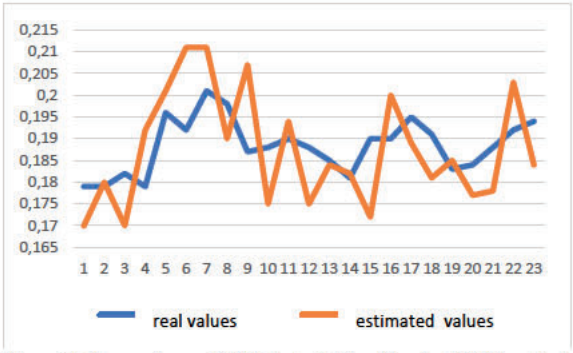


Figure 7: Comparison of XPR data obtained by the KNN method.

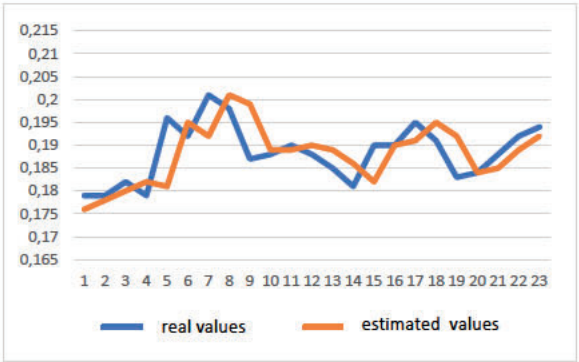


Figure 8: Comparison of XPR data obtained by the SVR method.

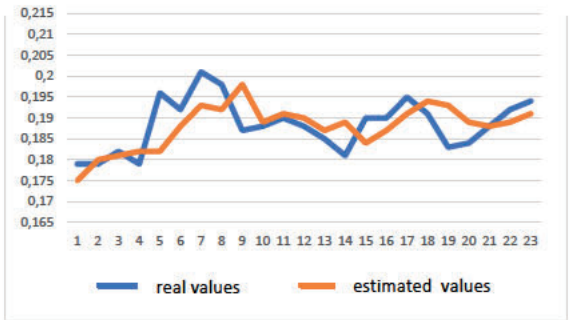


Figure 9: Comparison of XPR data obtained by the RF method.

When the graphs of the three methods are examined, it is seen that the most successful method in XRP price analysis is the

DVR. Although the DVR and RO methods gave very close OMH values, significant deviations were observed in the estimated data obtained by the RO method. Compared to the Random Forest algorithm, its response to price changes was much slower. When the graph of the KNN method is examined, it can be said that it gives unsuccessful results in all respects compared to the other two methods.

IV. Conclusion

Cryptocurrencies are digital assets that have different characteristics to achieve different purposes, many of which have a total supply. Crypto currencies are one of the investment tools that have attracted the attention of all kinds of investors since 2009, when they were created. Although the main purpose of crypto money is to be used in payment and money transfer transactions, it has also been a commodity used by a large mass as an investment tool. Although there are many different negative evaluations these days, when technological developments begin to change the financial system, they will be used as an investment tool in the future, given the value and potential of cryptocurrencies. So much so that although it is not supported by any state or institution, Bitcoin, the first and most popular of crypto currencies, can constantly attract the attention of new investors. For this reason, predicting the future price of crypto coins is among the indispensable for investors.

In this study, the price prediction was made with KNN, SVR and RF algorithms using the daily opening, closing, highest and lowest price, daily volume and daily total market value attributes of BTC, ETH and XRP, the three most valuable crypto currencies in the cryptocurrency world.

The price estimation is made on the basis of USD so that it is not affected by exchange rates. The data set to be used was created by extracting all the data from the first date when the relevant crypto money was traded on stock exchanges until today.

In the study, it was seen that some attributes in the data sets consist of missing data and all the missing data were deleted and the data sets were made ready for use. All data sets were normalized within themselves, then separated as training and prediction data set and made ready for analysis. With the WEKA 3.8 (Hall et al. 2009) program, the training data sets were trained with the most ideal parameters and k times cross validation values by changing the k-fold cross validation values with the parameters specific to K Nearest Neighbor, Support Vector Regression and Random Forest algorithms. 23-day price estimation was made using the models created. Price predictions have been compared with actual data. As a result of the comparison, the success ranking of the methods was SVR, RF and KNN, respectively. With the Support Vector Regression algorithm, the lowest average absolute error results were obtained in Bitcoin, Ethereum and XRP price prediction. Since the K Nearest Neighbor algorithm gives the highest absolute error results, it was determined as the most unsuccessful method among the three methods.

As a result, when the estimated data of the SVR algorithm are compared with the real data, it is seen that the estimated price data match the actual data and react very well to price changes compared to other methods. Considering the success of the Support Vector Regression method in this study, it has been concluded that it can be applied to all other cryptocurrencies in the exchanges where BTC, ETH and XRP are traded.

In future studies, a data set with more attributes can be used by adding some indicator values used in technical analysis in addition to the features used in this study. Again, in another study in the future, after estimating the features used in this study with different attributes, the estimation model can be analyzed in two stages by making a price estimation with estimated attribute values.

Acknowledgment

This study has been supported by Eskisehir Technical University Scientific Research Projects Committee (ESTUBAP-19ADP048).

References

[1] K Alkhatib, H Najadat, I Hmeidi, and M K A Shatnawi, "Stock price prediction using k-nearest neighbor (kNN) algorithm," *International Journal of Business, Humanities and Technology,* vol 3 (3), pp 32-44, 2013

[2] K İlarslan, "K-en yakın komşu (KNN) algoritması ile hisse senedi fiyatlarının tahmin edilmesi: Bist'den örnek bir uygulama," *Akademik Sosyal Araştırmalar Dergisi*, vol 30, pp 375-392, 2016

[3] E Filiz, H A Karaboğa, and S Akogul, "Bist- endeksi değişim değerlerinin sınıflandırılmasında makine öğrenmesi yöntemleri ve yapay sinir ağları kullanımı," *Çukurova Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, vol 26(1), pp 231-241, 2016

[4] S Lahmiri and S Bekiros, "Intelligent forecasting with machine learning trading systems in chaotic intraday Bitcoin market," Chaos Solitions & Fractals, Vol 133, pp 1-7, 2020

[5] E Fradinata, Z M Kesuma, S Rusdina, and N Zaman, "Forecast analysis of instant noodle demand using support vector regression (SVR)," in *Conf. Rec. 2019 1st South Aceh International Conf. on Engineering and Technology (Saicoet): IOP Conference Series-Materials Science and Engineering,* pp 506.

[6] G Meral, "Türkiye'deki enerji santrallerinde doğal gaz tüketiminin destek vektör regresyon ile tahmini," Yüksek Lisans Tezi, Fen Bilimleri Enstitüsü, Afyon Kocatepe Üniv , Afyon, 2019

[7] H Ince and T B Trafalis, "Short term forecasting with support vector machines and application to stock price prediction," *International Journal of General Systems*, vol 37(6), pp 677-687, 2008

[8] G P Herrera, M Constantino, B M Tabak, H Pistori, J J Su, and A Naranpanawa, "Long-term forecast of energy commodities price using machine learning," *Energy*, vol 179, pp 214-221, 2019

[9] B Baba and G Sevil, "Predicting IPO initial returns using random forest," *Borsa İstanbul Review,* vol 9(1), pp 13-23, 2020

[10] I Kumar, K Dogra, C Utreja, and P Yadav, "A comparative study of supervised machine learning algorithms for stock market trend prediction," in *Conf. Rec 2018 Int. Conf. on Inventive Communication and Computational Technologies (ICICCT),* pp 1003-1007

[11] D C A Mallqui and R A S Fernandes, "Predicting the direction, maximum, minimum, and closing prices of daily bitcoin exchange rate using machine learning techniques," *Applied Soft Computing Journal*, vol 75, pp 596-606, 2019

[12] Z Chen, C Li, and W Sun, "Bitcoin price prediction using machine learning: An approach to sample dimension engineering," *Journal of Computational and Applied Mathematics*, vol 365,pp 1-13, 2020

[13] M Hall, E Frank, G Holmes, B Pfahringer, P Reutemann, and I H Witten, "The WEKA data mining software: An update," *SIGKDD Explorations*, vol 11(1), pp 10-18, 2009

[14] V N Vapnik, *Statistical Learning Theory* New York: John Wiley and Sons, 1998

[15] A Cutler, D R Cutler, and J R Stevens, "Random Forests," *Machine Learning*, vol 45(1), pp 157- 176, 2011

Cloud Computing Based Voting Classifier Method Used for Survival Prediction of Heart Failure Patients

Ö. YAKUT

Kocaeli University, Kocaeli/Turkey, onder.yakut@kocaeli.edu.tr

Abstract **- Millions of people die every year due to functional impairments in different organs caused by Cardiovascular diseases (as myocardial infarctions and heart failures). By analyzing the medical records kept during the treatment of patients with heart failure, it can be ensured that clinicians pay attention to situations that are difficult to detect. In this study, using the medical records of the patients, it was predicted whether the patients would survive with the heart failure clinical records data set. In this study; Machine learning methods such as Artificial Neural Network (ANN), Logistic Regression (LR), Random Forest (RF), Naive Bayes (NB), Support Vector Machine (SVM) and k-Nearest Neighbors (kNN) have been used. Ensemble Learning-based Voting Classifier (VC) method which utilized the predictions of these machine learning models and makes decisions according to the majority vote has been proposed. The performance results of these models have been evaluated and compared. Performance values of the Voting Classifier method have been obtained as F1-Score 0.76%, Accuracy 0.83%, Precision 0.79%, Recall 0.75%, ROC 74.7% and ROC-AUC 0.86%. In this study, it has been concluded that the performance of the proposed Voting Classifier model is higher compared to the study in the literature. It has been concluded that the proposed method can be used to assist clinicians in their decision-making process to predict whether a patient will survive with heart failure. The machine learning methods developed and suggested in the study were coded with the Phyton programming language using Google Colaboratory (Colab notebook) which is Google Cloud Computing service.**

Keywords **- Cloud Computing, Data Science, Google Colaboratory, Machine Learning, Voting Classifier.**

I. Introduction

ACCORDING to the data of the World Health Organization (WHO) [1], approximately 17.7 million people die each year due to heart diseases and functional impairments in different organs caused by these diseases. Heart failure is the condition in which the heart cannot pump enough blood to the body with the occurrence of cardiac functional disorders. Patients with heart failure have a very high mortality rate. The frequency and prevalence of this disease are increasing worldwide. Early diagnosis of heart failure is extremely important. Today, there is a need for decision support systems that will assist clinicians in the diagnosis of diseases in the field of health. Decision support systems analyze the historical data of patients using machine learning methods and make predictions for the current situation.

Various researches have been conducted using machine learning methods to predict heart failure.

Angraal et al. [2] estimate the risk of mortality and hospitalization for heart failure in patients with heart failure with preserved ejection fraction using machine learning techniques. They proposed a decision-making tool to predict the course of the disease of these patients. Adler et al. [3] created a risk score that highly predicted mortality in patients with heart failure using a machine learning approach. Marcinkiewicz-Siemion et al [4] proposed a tool to diagnose heart failure with an untargeted metabolomics and reduced ejection fraction of machine learning algorithms resulting from the combination of the two methods. Mathis et al. [5] developed machine learning models that include logistic regression, random forest, and extreme gradient boosting using perioperative data, including preoperative, intraoperative, to potentially detect heart failure with reduced ejection fraction at an early stage. Jing et al. [6] proposed a strategy to manage populations of patients with heart failure using large clinical datasets and machine learning methods.

In this study, the Ensemble Learning-based Voting Classifier method using various machine learning methods has been proposed. With this method, survival estimation of patients with heart failure has been made. This method has been thought to provide support to clinicians in the decision-making process regarding heart failure. While developing the proposed method, the hardware and software needed have been provided by using Cloud Computing based Google Colaboratory (Colab) .

II. Materials and Methods

A. Data Set

This data set includes medical records of 299 heart failure patients collected during the follow-up period of the Faisalabad Institute of Cardiology and Allied Hospital in Faisalabad (Punjab, Pakistan) between April and December 2015 with each patient profile having 13 clinical features [7-10]. The 299 patients with heart failure consisted of 105 women and 194 men and were between 40 and 95 years of age. All of the patients had left ventricular systolic dysfunction and the heart failure stages of the patients were determined by the New York Heart Association (NYHA) as

class III or IV [8, 11].

Table 1: Heart failure clinical records data set feature ranking [8]

Clinical Feature	Feature Rank	Feature Group
Ejection fraction	1	DS1, DS2
Serum creatinine	2	DS1, DS2
Age	3	DS1
Creatinine phosphokinase	4	DS1
Sex	5	DS1
Platelets	6	DS1
High blood pressure	7	DS1
Anaemia	8	DS1
Smoking	9	DS1
Diabetes	10	DS1
Serum sodium	11	DS1

In Table 1 [8] Heart failure clinical records data set features are listed. The features in this list are in order of importance from higher to lower. Two separate feature sets are obtained from this feature set. The first of these is Data Set 1 (DS1), which consists of all features. The second is Data Set 2 (DS2), which consists of only the first two features. These two data sets are used for prediction.

B. Google Colaboratory

Google Colaboratory (Google Colab, Colab notebook) is a research project that offers powerful hardware options such as Graphics Processing Units (GPU) and Tensor Processing Units (TPU) to prototype machine learning methods. Google Colab offers a serverless Jupyter notebook environment for users to develop interactive code [12]. To use Google Colab, go to https://colab research.google.com. You can then log in with your existing Google account and access Google Colab's home page as shown in Figure 1. You can use the Google Colab service for free.

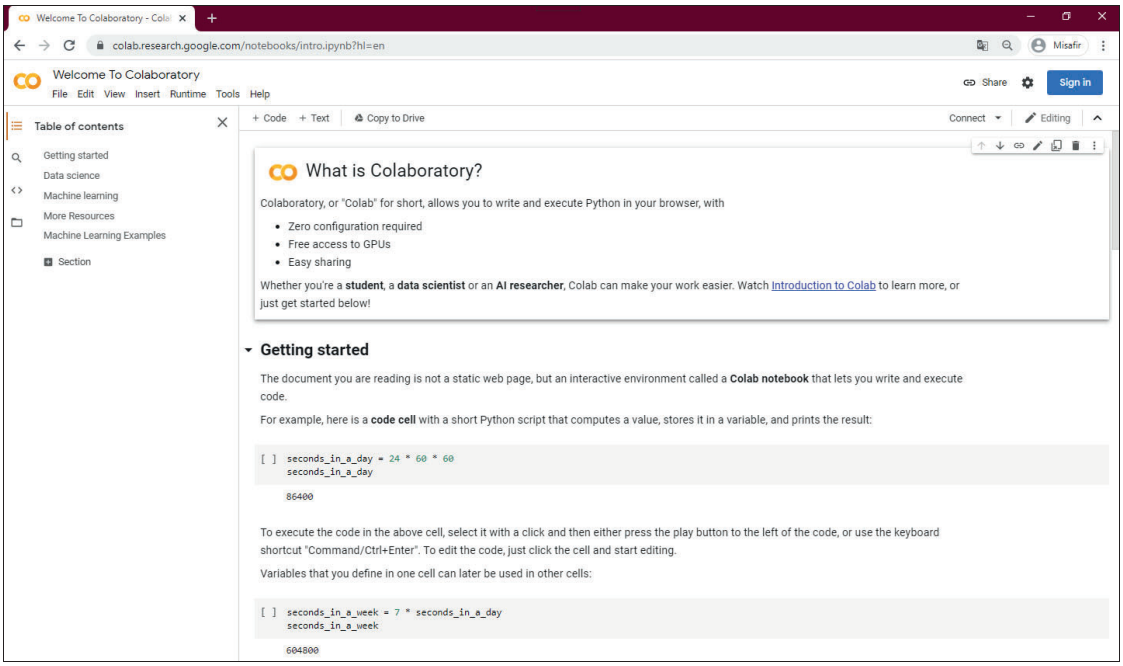


Figure 1: Screenshot of Google Colab homepage

C. Voting Classifier

In the Voting Classifier method, different types of machine learning methods are trained and tested on the data set to be studied. The output of these methods is then combined to get an ensemble estimate. The class that receives the majority vote in this unification process is considered as the estimate of the ensemble. It is called the hard voting classifier according to the majority vote. Collectively the estimation process usually produces results with a better degree of accuracy than the best classifier in the ensemble. With the Voting Classifier method,

the error in the ensemble estimation is reduced and thus the prediction success of the ensemble is increased.

In this study, six different types of machine learning methods such as Artificial Neural Network (ANN), Logistic Regression (LR), Random Forest (RF), Naive Bayes (NB), Support Vector Machine (SVM) and k-Nearest Neighbors (kNN) have been developed. The final estimation has been made by combining the results of these machine learning methods with the Voting Classifier (VC) method according to

the majority vote. This work has coded in Python programming language using Colab notebook. Scikit-learn [13] open-source machine learning library has been utilized for predictive data analysis . While developing machine learning methods each method has its own hyperparameters.

Hyperparameters of machine learning methods have been used with their default values in scikit-learn machine learning library.

D. Performance Metrics

Accuracy is a classification score. In multi-label classification, this function calculates subset accuracy. Accuracy score ranges from 0 to 1. F1 Score is interpreted as the weighted average of precision and recall values. The F-Score score reaches its best value at 1, and its worst value at 0. The recall is the classifier's ability to intuitively find all positive samples. Precision is the ability of the classifier to intuitively not label a sample positive that is negative. The best value of precision and recall is 1 and the worst value is 0 [14]. ROC shows the trade-off between precision and recall. It explains how well the model is at predicting positive class. The result of the estimation is the area calculation in Area Under the Receiver Operating Characteristic Curve (ROC AUC). ROC-AUC demonstrates the model's ability to distinguish classes. ROC and ROC-AUC values are 1 for the most perfect classifier and 0 for the worst classifier [14, 15].

III. Experimental Study

In this study, a cloud computing-based Voting Classifier method is proposed as in Figure 2 to predict whether patients with heart failure will survive using heart failure clinical data set. The user can access Google Drive and Google Colab services by logging into the system with a Google account. The user uploads the data set to Google Drive. Interactively, Colab can also run the codes of machine learning algorithms developed in Python on the Google Cloud Computing Platform and get the results. In this way, the user provides his software and hardware needs by saving time and money as he benefits from cloud computing services. Users can make the necessary analysis for themselves without dealing with the maintenance, repair, software update and licensing of the cloud computing platform they use. It can distribute these analysis results and their developed models on the web.

In the proposed method, features in the heart failure clinical data set have been normalized according to Min-Max normalization. Next, the Heart failure clinical dataset (299 records in total) has split into two datasets; the training data set is 209 records (2/3) and the test set is 90 (1/3) records.

The training and test data sets have been fed into machine learning models developed using the 5-fold cross-validation technique. Artificial Neural Network, Logistic Regression, Random Forest, Naive Bayes, Support Vector Machine and k-Nearest Neighbors machine learning methods used in this study have been developed. First, these methods have been trained with the training data set. Then, these methods have been tested with the test data set. Next, the prediction results of these models have been obtained. Finally, the prediction results of all these models have been predicted by the Voting Classifier machine learning method which decides according to the majority vote. Consequently, the final prediction has been made whether patients with heart failure will survive using the Voting Classifier machine learning method. Since the proposed method has been developed based on cloud computing it can be deployed on the internet.

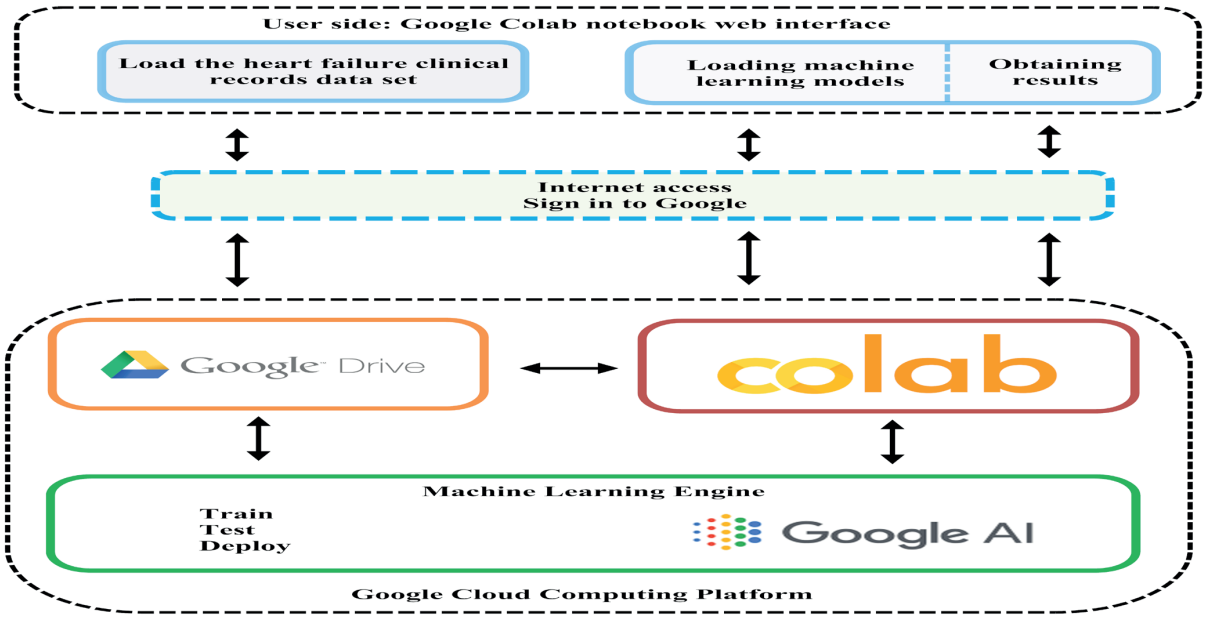


Figure 2: Block diagram of the proposed prediction method

IV. Experimental Study

The heart failure clinical data set is grouped in order of importance in terms of features [8]. As the first group, all the features (11) of the data set have been used and named as Data Set 1 (DS1). Ejection fraction (EF) and serum creatinine (SC) features have been used and named as Data Set 2 (DS2). The machine learning methods in this study have been trained and tested using these two data sets.

In this study, the machine learning methods used DS1 to predict whether patients with heart failure will survive. The performance results of machine learning methods as a result of this prediction are shown in Table 2. The methods in Table 2 are listed from high to low performance.

Table 2: Prediction results of machine learning methods trained and tested using DS1

Model		F1-score	Accuracy	Recall	Precision	ROC	ROC-AUC	Rank
Voting Classifier	Train	0,827	0,831	0,830	0,860	0,867	0,920	1
	Test	0,840	0,857	0,857	0,864	0,755	0,940	
Random Forest	Train	0,839	0,864	0,820	0,870	0,820	0,940	2
	Test	0,826	0,880	0,814	0,842	0,811	0,900	
Logistic Regression	Train	0,777	0,810	0,810	0,804	0,764	0,890	3
	Test	0,826	0,840	0,827	0,814	0,741	0,860	
Support Vector Machine	Train	0,863	0,860	0,850	0,879	0,851	0,950	4
	Test	0,802	0,760	0,800	0,817	0,775	0,840	
Naive Bayes	Train	0,743	0,740	0,740	0,737	0,678	0,860	5
	Test	0,761	0,780	0,780	0,759	0,663	0,830	
k-Nearest Neighbors	Train	0,830	0,838	0,840	0,847	0,788	0,910	6
	Test	0,752	0,760	0,760	0,749	0,674	0,730	
Artificial Neural Network	Train	0,961	0,960	0,960	0,963	0,960	1,000	7
	Test	0,749	0,730	0,750	0,700	0,746	0,800	

Table 3: Prediction results of machine learning methods trained and tested using DS2

Model		F1-score	Accuracy	Recall	Precision	ROC	ROC-AUC	Rank
Voting Classifier	Train	0,760	0,775	0,780	0,772	0,705	0,880	1
	Test	0,769	0,783	0,780	0,768	0,673	0,730	
Random Forest	Train	0,823	0,823	0,820	0,824	0,779	0,890	2
	Test	0,765	0,776	0,780	0,776	0,675	0,780	
Artificial Neural Network	Train	0,725	0,763	0,720	0,746	0,716	0,790	3
	Test	0,765	0,765	0,770	0,753	0,668	0,690	
Logistic Regression	Train	0,754	0,754	0,750	0,756	0,668	0,770	4
	Test	0,750	0,750	0,750	0,750	0,611	0,710	
k-Nearest Neighbors	Train	0,826	0,828	0,830	0,818	0,799	0,890	5
	Test	0,724	0,724	0,720	0,724	0,665	0,710	
Naive Bayes	Train	0,722	0,722	0,720	0,730	0,630	0,770	6
	Test	0,720	0,724	0,724	0,719	0,542	0,670	
Support Vector Machine	Train	0,754	0,763	0,760	0,752	0,714	0,790	7
	Test	0,719	0,719	0,720	0,719	0,584	0,690	

In this study, the machine learning methods used the DS2 data set to predict whether patients with heart failure will survive. As a result of the predictions, the performance results of machine learning methods are given in Table 3. The methods in Table 3 are listed from the high performance to the lower one.

V. Discussion

In this study, the performance results in Table 2 and Table 3 were obtained in order to predict whether patients with heart failure will survive. When the performance results in these Tables are examined, it is seen that all models have sufficiently high degree estimates. It is understood that the highest performance result in both Tables belongs to the

Voting Classifier machine learning method. This performance results show that the proposed method is more predictive in decision making.

When the results in Table 2 of the DS1 data set are examined, the performance of the methods used is high. The Voting Classifier method suggested in this study gave the highest performance result. The proposed method also produces less erroneous results as it makes the final decision using the predictions of other methods. Thus, instead of the prediction result of a machine learning method, the prediction result of more methods was evaluated. So, it has been seen that the ensemble prediction result is better than the prediction made alone.

When the results of the DS2 data set were examined, the results in Table 3 were obtained by using only two features. The performances of the obtained results are acceptably high. The proposed Voting Classifier method has the highest performance result.

Table 4: Comparison of prediction results

Model	F1-score	Accuracy	Recall	Precision	ROC	ROC-AUC
Logistic Regression [8] (DS1)	0,714	0,833	0,780	0,856	0,612	0,818
Voting Classifier (proposed method) (DS1)	0,840	0,857	0,857	0,864	0,755	0,940
Logistic Regression [8] (DS2)	0,719	0,838	0,785	0,860	0,617	0,822
Voting Classifier (proposed method) (DS2)	0,769	0,783	0,780	0,768	0,673	0,730

When the results given for the DS1 data set in Table 4 were examined, the proposed method gave a higher prediction result than the other method. When the results given for the DS2 data set in Table 4 are analyzed, the proposed method gave prediction results in the same range and very close to each other with the other method.

The proposed method obtained sufficiently high prediction results for both data sets (DS1 and DS2). Thus, it was concluded that the proposed method to predict whether patients with heart failure will survive will provide support to clinicians in the decision-making process.

VI. Conclusion

In this study, the Ensemble Learning-based Voting Classifier method has been proposed. With the proposed method survival prediction of patients with heart failure has been made. Sufficiently high predictions have been obtained using the proposed method. Thus, it has been revealed that the proposed method to predict whether patients with heart failure will survive will provide support to clinicians in the decision-making phase. The hardware and software requirement of the proposed method has been met by using Cloud Computing based Google Colaboratory (Colab). In future work, the performance of prediction results can be further improved by using more different machine learning methods and more comprehensive data sets.

References

[1] World Health Organization Cardiovascular diseases (CVDs), Available: http://www who int/mediacentre/factsheets/fs317/en/, [Accessed 19 10 2020]

[2] Angraal S , Mortazavi B J , Gupta A , Khera R , Ahmad T , Desai N R , Krumholz H M , "Machine learning prediction of mortality and hospitalization in heart failure with preserved ejection fraction", *JACC: Heart Failure*, vol 8(1), pp 12-21, January 2020

[3] Adler E D , Voors A A , Klein L , Macheret F , Braun O O , Urey M A , Greenberg B , "Improving risk prediction in heart failure using machine learning", *European Journal of Heart Failure*, vol 22(1), pp 139-147, January 2020

[4] Marcinkiewicz-Siemion M , Kaminski M , Ciborowski M , Ptaszynska-Kopczynska K , Szpakowicz A , Lisowska A , Kaminski K A , "Machine-learning facilitates selection of a novel diagnostic panel of metabolites for the detection of heart failure", *Scientific Reports*, vol 10(1), pp 1-11, January 2020

[5] Mathis M R , Engoren M C , Joo H , Maile M D , Aaronson K D , Burns M L , Najarian K , "Early detection of heart failure with reduced ejection fraction using perioperative data among noncardiac surgical patients: a machine-learning approach", *Anesthesia & Analgesia*, vol 130(5), pp 1188-1200, May 2020

[6] Jing L , Cerna A E U , Good C W , Sauers N M , Schneider G , Hartzel D N , Stough J V , "A Machine Learning Approach to Management of Heart Failure Populations", *JACC: Heart Failure*, vol 8(7), pp 578-587, July 2020

[7] Heart failure clinical records Data Set, Available: https://archive ics uci edu /ml /datasets/Heart+failure+clinical+records, [Accessed 07 10 2020]

[8] Chicco D, & Jurman G , "Machine learning can predict survival of patients with heart failure from serum creatinine and ejection fraction alone", *BMC medical informatics and decision making*, vol 20(1), pp 16, February 2020

[9] Ahmad T , Munir A , Bhatti S H , Aftab M , Raza M A , "Survival analysis of heart failure patients: A case study", *PloS one*, vol 12(7), pp 1-8, July 2017

[10] Ahmad T, Munir A, Bhatti SH, Aftab M, Ali Raza M , "Survival analysis of heart failure patients: a case study Dataset", Available: https: //plos figshare com/articles/Survival_analysis_of_heart_failure_patients _A_case_study/5227684/1 [Accessed 23 10 2020]

[11] Bredy C , Ministeri M , Kempn, A , Alonso-Gonzalez R , Swan L , Uebing A , Dimopoulos K , "New York Heart Association (NYHA) classification in adults with congenital heart disease: relation to objective measures of exercise and outcome", *European Heart Journal-Quality of Care and Clinical Outcomes*, vol 4(1), pp 51-58, January 2018

[12] Bisong Ekaba, *Building Machine Learning and Deep Learning Models on Google Cloud Platform*, Apress, 2019

[13] Scikit-learn Machine Learning in Python, Available: https://scikit-learn org/, [Accessed 27 10 2020]

[14] Sklearn Metrics, Available: https://scikit-learn org/stable/modules/classes html? highlight=metrics#module-sklearn metrics, [Accessed 27 10 2020]

[15] ROC Curves and Precision-Recall Curves for Imbalanced Classification, Available: https://machinelearningmastery com/roc-curves-and-precision-recall-curves-for-imbalanced-classification/, [Accessed 27 10 2020]

DETECTION OF AUTOMOBILE FAILURES BY SOUND PROCESSING AND DRIVER GUIDANCE SYSTEM DESIGN

Ali QUTUB[1] and İsmail SARITAS[1]

[1] Selcuk University, Konya/Turkey, akutub87@gmail.com
[2]Selcuk University, Konya/Turkey, ismailsaritas@gmail.com

Abstract – **Voice recognition and processing; in today's world, realistic authentication has become a leading technique in the matters such as detection of faults, access to control techniques, and finding solutions to problems. The presence of malfunctions in vehicles leads to difficulties in system integrity. Car-key, engine, wrong pin, wrong codes or other authentication errors, violation can be detected by stolen or hacked/lost or by transferring data to computer. In addition, these unsolicited statuses may cause access to this protected system to be stopped. However, voice segmentation can provide a complete solution to this kind of dilemma through automotive technologies. The system, which is applied as voice segmentation, consists of two steps: recognition and verification. The recognition is defined as multiple, i.e., comparison of a recorded data for a database with the whole system, while verification is defined as a one-to-one, i.e. comparison of a pair to determine whether they are caused by the same error.**

Keywords – **Artificial Intelligence, Embedded System, Voice Recognition, Voice Processing, Car Fault Detection.**

I. INTRODUCTION

In this study, it is aimed to detect many malfunctions such as mechanical, engine, electrical and electronic malfunctions that may occur in vehicles that have almost completely turned into electronic control systems in the automotive industry by using sound characteristics. An intelligent interface design will also be implemented that will guide the driver depending on the severity of the detected fault and the condition to be remedied. Thus, an early diagnosis of a malfunction that may occur in vehicles will be provided and larger failures that may arise from these malfunctions / failures will be prevented. In addition, by preventing accidents caused by malfunction, both loss of life and economic losses can be prevented. In this thesis project, firstly, sound recordings of real vehicles will be taken in defective and intact conditions, and data sets will be created by signal processing. What the real fault is, its importance and what needs to be done will be entered into the database as control data by an expert. The system tested in the results obtained will also be tested in vehicles and statistically analyzed. With an intelligent interface, the driver will be informed of what the fault (s) are as well as what to do.

Since 1990, DRA (Direction record assembly) has been researching speech recognition technology applications for speech and language development for young children as a speech research unit. The first project, known as STAR (Speech Training Aid Research), was supported by the awareness that teachers needed a "speech training aid" tool to assist young children in developing various communication and language skills[1]. They performed hardware speech recognition in their studies. They used RSC 364 card and PIC 16F876 for this operation. They carried out different experiments, both inside and outside the laboratory[2]. in his work; They stated that embedded systems are at the heart of all electronic devices of today, from mobile phones to garage door openers and medical devices. It is almost impossible to produce an electronic device without hardware and software[3]. in his work; They stated that the difference between embedded systems and computers is that computers are not designed to perform a single special function, but to perform many functions. The term general purpose computer is used to reinforce this difference[4]. Usually, an embedded system is an element of a large system. For example, modern cars have different embedded systems that can perform many different, special functions. One embedded system controls antilock brakes, the other monitors and controls the vehicle's emissions, while another displays these data on the dashboard[5]. Sound is the mechanical vibrations created during the transfer of energy in a solid, liquid or gas environment. Sound can also be defined as the response of humans to the pressure changes caused by the vibration of objects in a liquid, solid or gas environment. The vibration of the object is required to generate sound. Waves caused by the vibration of objects are called sound waves[6]. The number of repetitions in a signal produced per unit time can be expressed as frequency, and the frequency can also be called the frequency space. Frequency unit is Hertz (Hz). Frequency determines the pitch and fineness of the sound. If the number of repetitions of the signal per unit time increases, that is, if the frequency increases, the sound becomes thinner, if the frequency decreases, the sound becomes thicker. In other words, frequency is the number of displacements or oscillations of a particle per unit time[7].

II. MATERIALS AND METHODS

A. Dataset

The database of this website has been created from 6 vehicles so far. The data are all taken from KONFIAT's Technical service department, which is shown Figure (1.1).

Figure (1.1) KONFİAT

B. Methods

- Angular

Angular; It is a TypeScript-based and open source framework used to build web, mobile, and desktop applications. Angular applications include the use of Typescript, a language built on Javascript, together with Html, Css. The code written in Typescript is compiled into Javascript first and then processed in the browser. Using Typescript is not mandatory for Angular application development, but its ease of use and advantages make it more useful to use.

- Typescript

TypeScript is a strongly-typed, object-oriented and compiled open source programming language. Designed by Anders Hejlsberg, a Microsoft employee and designer of the C # language, the first version was published in 2012. It is possible to define TypeScript as both a language and a set of tools. It can be used to write JavaScript programs that can run in a client or server environment.

- Javascript

JavaScript, often abbreviated as JS, is a multi-paradigm, prototype-based, mandatory and functional scripting language commonly used in web browsers. Thanks to client-side scripts written in JavaScript, functions such as interacting with the user, controlling the browser, communicating with the server asynchronously and changing the web page content are provided. JavaScript is also widely used on the server side, thanks to platforms such as Node.js.

- HTML

Hypertext Markup Language (HTML) is the standard text markup language used to create web pages, and is a system of pieces of code called tags that determine how documents will link to each other and how text and images within a document will be placed. Hypertexts used to link documents with each other are very powerful. The World Wide Web (shortly WWW or Web), which lives on the Internet, is used to create interactive, multi-platform, multimedia and client / server applications behind the HTML system.

- Cascading Style Sheets (CSS)

Cascading Style Sheets (CSS), called "Cascading Style Sheets" or "Cascading Style Templates", is a markup language that offers extra possibilities in the field of text and formatting in addition to HTML. CSS, in its simplest form, allows us to visually improve our HTML tags (size, color, font, etc.).

There are 3 style page types: Internal, External and In-line. Internal is the way we write our CSS code in the style tag at the top of the page. External (external) is the writing format that we open a document with .css extension for our CSS codes and write here, then call this document into HTML with the link tag. In-line is the format in which we write our CSS codes in-line into our HTML structure.

- Firebase

Firebase is a cloud-based platform developed by Google for building mobile and web applications where we can store our JSON data in real time on a remote server. Firebase, which handles applications such as application management, usage tracking, data storage, sending notification without the need to write an extra server and server-side code, provides separate access for each application equipped with features such as Realtime Database, Notification, Remote Config in its new developer-friendly interface.

- Amazon Web Services

It is a subsidiary of Amazon that provides on-demand cloud computing platforms and application programming interfaces. It was established in July 2002 and provides cloud computing web services, various basic abstract technical infrastructure and distributed computing building blocks and tools to individuals, companies and governments on a pay-as-you-go basis. The virtual computer version of AWS emulates many of the attributes of a real computer, including hardware central processing units and graphics processing units for rendering.

- AWS S3

Amazon S3 is short for Amazon Storage Service in AWS service. AWS S3 provides IT teams, Developers, and highly secure, durable storage with low latency in the AWS cloud. It is basically an object storage service of AWS that can store and receive unlimited data from anywhere in the world.

S3 provides 99.999999999% durability for objects stored in the service and supports multiple security and compliance certificates. The administrator can also connect S3 to other AWS security and monitoring services such as CloudTrail, CloudWatch, and Macie. There is also an extensive partner network that connects their services directly to S3.

Amazon S3 is a large-scale storage service with millions of cloud users worldwide. Amazon S3 is designed with a focus on simplicity and robustness with very few feature sets. In addition, it has architectural features suitable for the power and scaling nature of cloud computing.

- AWS Amazon SageMaker

Amazon SageMaker is a cloud machine learning platform released in November 2017. Amazon SageMaker is a fully managed service that provides any developer and data scientist with the ability to rapidly develop, train, and deploy machine learning (ML) models. SageMaker makes it easier to develop high-quality models by taking on the heavy loads in every step of the machine learning process. SageMaker also allows developers to deploy ML models to embedded systems and edge devices.

- AWS DynamoDB

Amazon DynamoDB is a key-value and document NoSQL database provided by Amazon Web Services that delivers

single-digit performance in milliseconds at any scale. NoSQL; Alternative to relational databases, it is a database solution that offers high performance and ease of management in systems with non-relational, flexible structure, big data and many active users.

- AWS Amazon Elastic Compute Cloud ec2

Amazon Elastic Compute Cloud (EC2) is a web service from Amazon Web Services that provides secure, resizable compute capacity in the cloud. The platform was established on August 25, 2006 and is designed to facilitate web-scale cloud computing for Developers. Amazon EC2 allows users to lease virtual computers where they can run their own computer applications. EC2 encourages scalable deployment of applications by providing a web service where a user can boot an Amazon Machine Image (AMI) to configure a virtual machine that Amazon calls an "instance" containing any software they want. A user can create, start, and terminate server instances as required, paying as much as seconds for active servers. EC2 provides latency optimization and high redundancy by giving users control over the geographic location of the cloud servers.

- AWS lambda

Applications developed with serverless architecture are event-based. In other words, an event must occur in order for our code to work. (HTTP request, cronjob, S3 file upload…) As a result of this event, our codes perform the determined operations.
AWS Lambda is one of the first services that comes to mind when we say 'Serverless'. In these services, we only write the code we want to run and make the necessary resource settings. Since we are not dealing with settings such as which server our code runs on, there is no need for settings such as management or maintenance, security.

- Filter banks

In signal processing, a filterbank is a band-pass filter array that separates the input signal into multiple components, each carrying a single frequency subband of the original signal. Another application of filter sets is signal compression when some frequencies are more important than others.

- Python

Python is an object-oriented, interpretive, modular and interactive high-level programming language. Python was started to be developed by Guido van Rossum in Amsterdam in the early 1990s. Contrary to popular belief, it takes its name not from the python snake, but from the show performed by Rossum's favorite comedy group called MonthyPython.
Simple syntax based on indents makes the language easy to learn and remember. This makes it a language where programming can be started without wasting time with the details of its syntax.
Python is an interpretive language, which means it can be run without compiling, unlike languages like C and C ++. This makes it easier to develop programs with Python. Many things, data structures, functions that you will need while writing a program in Python are presented to you. In this way, you can write programs in a much more rapid manner with the presentations infrastructure, without having to design down to the finest details to solve a problem as in other languages.
Python has a simple syntax. In this way, it becomes easier and more enjoyable to write programs, and programs written by others can be understood more easily. Python allows you to do a lot with little code.

C. *Why we used those teknolojis?*

(angular, html, css, js, typcript): We use these technologies to make a one-page application that allows us to use functions such as live recording, and we can also reuse the same code and technology in the mobile application.

(Firebase) We use it for user authentication (login and registration) from many websites such as Google, Apple and Facebook.
(AWS S3) Use S3 to store all of our engine sounds uploaded or uploaded by the user.
(AWS Amazon SageMaker) We use sagemaker to train our model in the cloud, this option is better in terms of time and performance.
(AWS dynamodb) We use dynamodb as our main database to store all information such as the location of sound in S3 and information about our user ... etc.
(Ec2) Use it as the main server for the backend and font application you can use to run our entire application. Use EC2.
(Lambda) Amazon offers a service lambda that will run functions in the cloud, separate from your server, to give us the security and scalability we need to manage users of all sizes.
(Python) It is the language in which we write our model and algorithm.

III. CONCLUSION

1.st car
Records were taken in the technical service department of KONFIAT firm in the experimental study. Until now, 2 vehicle data have been taken, and since the normal working voice records of the third of them cannot be taken due to Korana virus, data of 2 vehicles have been studied. While the vehicle is running, audio data is recorded on the phone. The recorded data is uploaded to the computer and analyzed. Recorded sound data is as in figure (3.1).

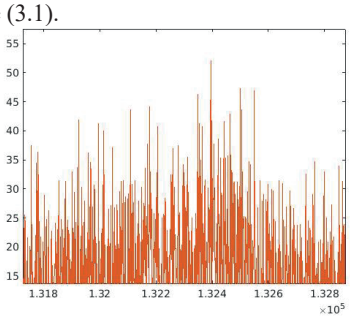


Figure 3.1: Defective sound wave

After the defective sound recording of the vehicle is uploaded to the system, filter banks are applied and after the sound wave is filtered, it is obtained as shown in figure (3.2).

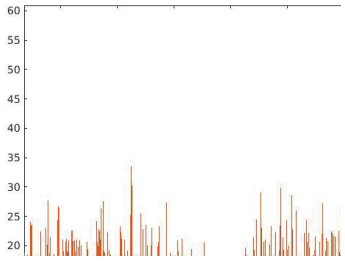


Figure 3.2: Sound wave filters

The vehicle is operated by KONFİAT after the malfunction is corrected and sound recording is made again in the same condition and place. The sound data obtained is as in Figure (3.3).

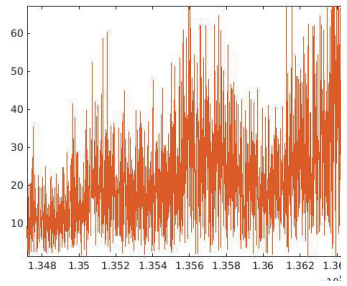


Figure 3.3: Normal sound wave

2.nd car

With the data obtained, the engines of the vehicles are loaded into the system and the sound wave appears in the system as in figure (3.4).

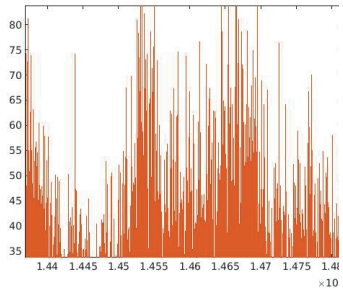


Figure 3.4: Defective sound wave

After the defective sounds of the vehicles are loaded into the system, filter banks are applied and after the sound wave is filtered, it looks like the figure (3.5).

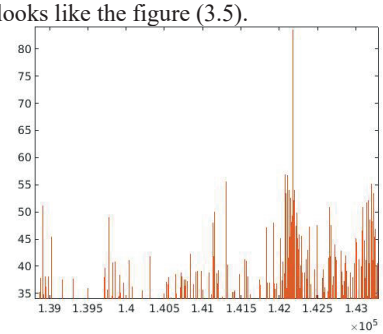


Figure 3.5: Sound wave filters

After the vehicle is in its normal state by CONFIAT, the sound wave looks like figure (3.6).

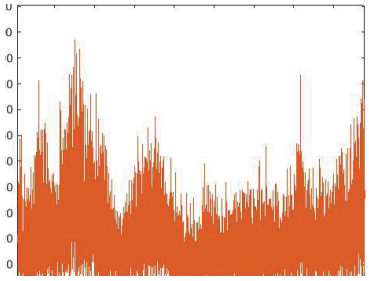


Figure 3.6: Normal sound wave

REFERENCES

1. Barr, M. ve Massa, A., 2006, Programming embedded systems: with C and GNU development tools, " O'Reilly Media, Inc.", p.43-51.
2. Benjamin, D. J., Berger, J. O., Johannesson, M., Nosek, B. A., Wagenmakers, E.-J., Berk, R., Bollen, K. A., Brembs, B., Brown, L. ve Camerer, C., 2018, Redefine statistical significance, *Nature Human Behaviour*, 2 (1), 6.
3. Bourlard, H. A. ve Morgan, N., 2012, Connectionist speech recognition: a hybrid approach, Springer Science & Business Media, p.192
4. Bressert, E., 2012, SciPy and NumPy: an overview for developers, " O'Reilly Media, Inc.", p.27-37
5. Çelebi, M., 2012, Ses komut tanıma ile gezgin araç kontrolü.
6. Davis, K. H., Biddulph, R. ve Balashek, S., 1952, Automatic recognition of spoken digits, *The Journal of the Acoustical Society of America*, 24 (6), 637-642.
7. Demirkale, S. Y., 2007, Çevre ve yapı akustiği: mimarlar ve mühendisler için el kitabı, Birsen Yayınevi, p.332-358
8. Engin, A., 2008, SAYISAL MODÜLASYON TANIMA SİSTEMLERİ İÇİN BAYES KARAR KURALLARI SINIFLANDIRICISININ KULLANIMI, *Engineering Sciences*, 3 (1), 117-123.
9. Femi, A., 2013, Perception of performance appraisal and workers' performance in Wema Bank headquarters, Lagos, *Global Journal of Arts, Humanities and Social Sciences*, 1 (4), 89-101.
10. Fezari, M. ve Bousbia-Salah, M., 2006, A voice command system for autonomous robots guidance, *9th IEEE International Workshop on Advanced Motion Control, 2006.*, 261-265.
11. Hinton, G., Deng, L., Yu, D., Dahl, G., Mohamed, A.-r., Jaitly, N., Senior, A., Vanhoucke, V., Nguyen, P. ve Kingsbury, B., 2012, Deep neural networks for acoustic modeling in speech recognition, *IEEE Signal processing magazine*, 29.
12. Kamal, R., 2011, Embedded systems: architecture, programming and design, Tata McGraw-Hill Education, p.3-1

Developing A Web Application with Django Architecture: An Example of Student Information System

C. AKKAN[1] and S. TASDEMIR[1]

[1] Selcuk University, Department of Computer Engineering, Konya Turkey, cihat.akkan@gmail.com
[1] Selcuk University, Department of Computer Engineering, Konya Turkey, stasdemir@selcuk.edu.tr

Abstract **- Software developers who want to develop web applications need auxiliary software to make their work easier. At this point, Django architecture built on Python language can be used. Django is an open source web interface that uses the MVT (Model, View, Template) architectural structure. Complex web applications can be developed easily with the advantages of Django architecture. It provides convenience to users in terms of modularity, dynamism, flexibility, security, detailed error reports, database management, management panel usage and working on every platform. In this study, using Django architecture built on Python language, a web application was made targeting an educational institution at high school level and the student information system was successfully concluded by running the designed web application.**

Keywords **- Django, Web Application, Python, Student Information System**

I. INTRODUCTION

With the rapid advancement of technology, the emergence of new languages and the increase in the number of user-friendly interfaces, libraries and modules integrated into these languages, extensive resources have emerged in the software world.

Amateur or professional users try to develop individual, corporate websites by using these resources. One of the most important points that users meet is that their projects are interrupted or incomplete as a result of not having a command of software languages. Hence, they desire to meet the needs such as doing a lot of work by writing less code with low software knowledge, using ready-made designs, developing user operations, taking security measures, developing dynamic web sites. In the absence of the necessary software competence, they need auxiliary tools that will enable them to do all these significant tasks on their behalf. Professional users use these auxiliary tools to facilitate their work, to automate, to provide the standard and to save time.

In the study, Django Web Framework, which has been developed on Python language, which has been introduced to the software world for web site development in recent years, will be discussed. Once we consider the Python language, we realize that Python is a programming language introduced by Guido Van Rossum in 1991 [1]. It is a new software that is easy to learn and coding of Python language compared to other languages, backs up object-oriented programming, is popular, its libraries are advanced and at a high level, the number of developers is high and it is used in scientific research (machine learning, deep learning, artificial intelligence, etc.) are among the reasons why it is preferred to develop [2]. Besides, Rossum's starting work at Google caused the Google company to use Python language a lot. The active use of the Python language by this company which is world-famous has greatly boomed the popularity of this language [3]. The Python language's ranking in the software world between 2016 and 2017 is shown in Figure 1. Spacepy, Dropbox, BitTorrent, NumPy, Matplotlib can be given as examples of popular applications and modules written in Python language [4].

May 2017	May 2016	Change	Programming Language	Ratings	Change
1	1		Java	14.639%	-6.32%
2	2		C	7.002%	-6.22%
3	3		C++	4.751%	-1.95%
4	5	^	Python	3.548%	-0.24%
5	4	v	C#	3.457%	-1.02%
6	10	^^	Visual Basic .NET	3.391%	+1.07%
7	7		JavaScript	3.071%	+0.73%
8	12	^^	Assembly language	2.859%	+0.98%
9	6	v	PHP	2.693%	-0.30%
10	9	v	Perl	2.602%	+0.28%

Figure 1: Most used programming languages [5]

When considering the Django framework, we realize that it is a web development interface built on the Python language and is a free and open source web application development structure. With Django, it enables you to overcome the difficulties met while developing web applications and to spend their energy and resources on the application they developed without wasting time with the basic problems. One of its main purposes is to ensure that developers can reach the completion of the application in the shortest time and easily from the concept [6].

In Kasapoğlu's work, he designed his web application with Django Web Framework while planning the resources used in software tests and software testing strategy, running software tests and keeping records related to tests, creating software test results reports, minimizing human errors in processes and performing repetitive work by the tool, standardizing processes and facilitating the documentation process [7].

Barışçı and Javadı discussed whether people's emotions and ideas about incidents can be determined from the content of

their posts on Twitter. In this study, Django Web Framework was used to develop the web application while selecting the Python programming language [8].

Aslan and Bilgin developed an interventional web application in cooperation with Python and Django Web Framework so as to embody the engineering intelligent control course, to absorb the course and to exercise more. Developed interactive applications are grouped under three headings: Routh table creation, Time response simulations and Control Action and simulations. It is estimated that the tool which is developed will contribute to control training [9].

Thanks to the student information system, they designed, Karaş and Baz ensured that educational activities were collected, stored digitally, queried on the internet and presented to users in documents and reports when needed [10].

In their studies, Ergin and Akseki examined the student information systems used in graduate education and they concluded that the workload was considerably minimized by checking the information of students, courses and academicians via the system. The systems they examined include Ankara and Gazi University Institute of Educational Sciences, Bahçeşehir University Institute of Social Sciences and Cambridge University Graduate Student Information and Registration systems [11].

Considering the necessity of the student information system to be developed with Python language and Django architecture, in our primary and secondary schools throughout our country, there is a need for information systems in order to record the social, behavioral and cognitive information of the students, to follow this information during the semester, and to use it in interdisciplinary relations. It is very significant that the information obtained can be stored, accessed when requested, and data security. The knowledge of the students in their past education levels will shed light on the future and concrete data will be used at the point of orientation. In order to eliminate this deficiency, it is aimed to design online systems and to use these systems by administrators and teachers. In this student information system, a vocational high school profile is taken into consideration and a student information registration system has been established, where teachers can become a member, show their thoughts about students as positive or negative, and the views of the relevant class or the relevant student can be accessed only when logging on.

II. Materials and Methods

Django is a high-level web framework designed on Python and licensed under a BSD license. It is different from other web framework infrastructures with its easy to use and easy installation, offering very detailed error reports and brand new interfaces. Its name comes from the jazz guitarist Django Reinhardt. Django project is to facilitate the use of complex web applications. Django is designed in a structure with reusability, modularity, rapid development process policy. Django also serves a dynamic management panel that includes basic record adding, deleting, editing and updating functions for each project. It is a comfortable and useful web framework for those eager to get rid of complex databases [12].

Django contains many extra features (user authentication, content management, sitemaps, RSS (Really Simple Syndication) feeds, etc.) to handle common web development tasks. In security matters, many security errors such as SQL (Structured Query Language) injection, cross-site scripting, cross-site request forgery and clickjacking help developers avoid. It provides a secure way to manage users' accounts and passwords with a user authentication system. Besides, Django serves the ability to rapidly and flexibly scale the peak traffic demands [6]. Examples of websites using Django are Disqus, Instagram, Knight Foundation, MacArthur Foundation, Mozilla, National Geographic, Open Knowledge Foundation, Pinterest, Open Stack [12]. The mentioned websites use Django's MVT (Model, View, Template) architecture shown in Figure 2.

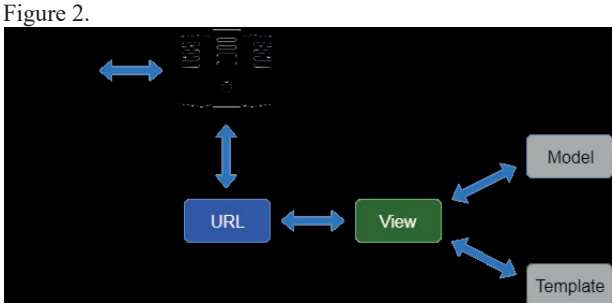


Figure 2: MVT architecture of Django [10].

The model statement includes the database structure and functions. While creating the model structure, the "class" logic in the programming structure is used. Each class actually corresponds to a table in the database. A part of the model layer belonging to the application developed in Figure 3 and Figure 4 is shown. The department and class tables of the Student Information System were created with the help of the model layer, and the information taken from the models was ordered by field with the expression "ordering". As can be seen in Figure 4, the class model is connected to the Section model with the help of foreign key.

```
    def __str__(self):
        return self.bolum_ad

    class Meta:
        ordering = ['bolum_ad']
        verbose_name_plural = "Bölümler"


class Sinif(models.Model):
    sinifin_bolumu = models.ForeignKey(Bolum, on_delete=models.CASCADE)
    sinif = models.CharField(max_length=20, verbose_name='Sinif')

    def __str__(self):
        return self.sinif

    class Meta:
        ordering = ['sinif']
        verbose_name_plural="Siniflar"
```

Figure 3: Model layer codes

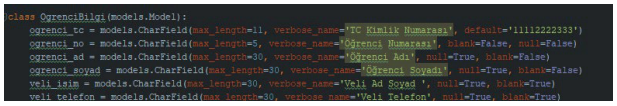


Figure 4: Model layer codes

In the View layer of MVT architecture, a bridge is set between the template and model layers. The requests from the user are processed and the responses are sent. By defining classes, functions and variables in the View layer, all kinds of programmatic operations and database listing, deletion, addition, update can be realized. The model layer can be thought of as the background skeleton of the MVT architecture, and the View layer as the brain that regulates the functioning of the architecture. In Figure 5, Figure 6 and Figure 7, a part of the View layer belonging to the application is shown.

Figure 5: View layer codes

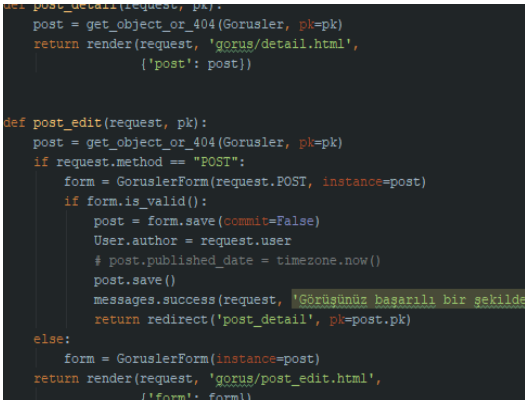


Figure 6: View layer codes

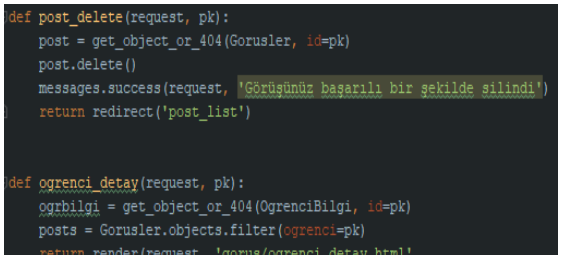


Figure 7: View layer codes

Because this layer contains visual design, user interaction and information taken from the database. Information from the View layer is presented to the user via the template language of Django. Again, one of the advantages of this layer is that it can run external functions with the "templatetags" structure offered by Django, so that this layer can call some information needed without going to the View layer. In Figure 8 and Figure 9, examples from the Template layer of the application are presented.

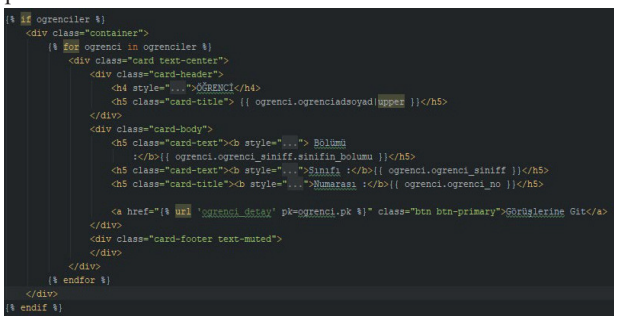


Figure 8: Template layer codes

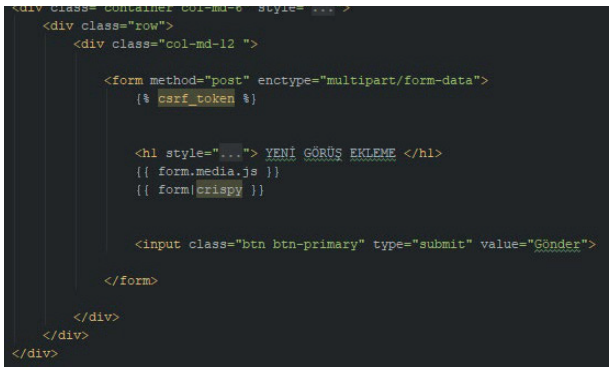


Figure 9: Template layer codes

One of the strongest features of Django is its Admin panel structure, which makes the user's work vitally easier, as well as its MVT architecture. In this way, the created models and applications are added to the management panel and the management of the web application becomes very easy. Introducing models and applications to the management panel is needed. Adding the model layers created above to the admin panel will make things easier in the future. Figure 9 shows

adding models to the management panel and the columns, search columns, filtering features of the table to be shown in this panel.

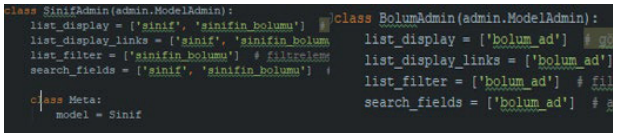


Figure 10: Codes for adding models to the admin panel

One of the basic needs while realizing the project is that the applications created with Django work in harmony on all kinds of devices. As the screen sizes of phones, tablets or computers will be different, the Bootstrap library was used in the study to provide these adjustments. Bootstrap is an open source library that allows the site to be viewed differently and proportionally to the size of your device for phone, tablet and desktop computers [13].

With the web application developed stemmed from this study, the teachers in charge can register to the system, update their membership information, examine the opinions of the classes, examine the students individually, add comments to all students registered in the school when necessary, update or delete the opinions they have added, and search for a student whose name is only remembered. Only the logged in user at the point of view will be able to realize this operation. In order for the system to be used comfortably in any environment, the responsive (flexible) design has been taken into consideration and the web application has been provided to work in harmony with all kinds of devices. All these systems are controlled with the benefits of Django architecture and Responsive design is provided with Bootstrap infrastructure. The visuals of the web application are shown in the following figures.

Figure 11: First two of the last views entered into the system

As seen in Figure 11, the opinions made can be filtered according to the teacher, class, subject and date. The filtering information is transmitted to the View layer and interpreted there and transmitted to the Template layer. If the opinions belong to the relevant user, the opportunity to edit them with the session control algorithms provided by Django has been given.

Figure 12: New comment adding screen

When the user logs in, the new comment adding screen becomes active and the students of the relevant class are brought automatically with the help of Django's Smart Selects module (Figure 12).

Figure 13: Detail screen of a created view

The detail page of the previously created view is shown in Figure 13. In this way, the user who wrote the view is permitted to edit or delete the view. Here, the User module of the Django infrastructure has been used. In order to be understood better by the users, if the opinion is positive, it is shaded with green, and if it is negative, it is shaded with red.

III. Recommendations

With the Django architecture used in this study, the management panel was created automatically, and critical issues such as database, security and membership processes were easily realized.

This situation presented that the targets were succeeded and the efficiency of the work was revealed. When the project, which was run in the local computer environment, was wanted to be installed on real servers, it was revealed that not every company providing server service in the market had Django support. This problem is expected to disappear with the spread of Django architecture over time. Another problem met is that Django projects' initial server setup process is more complex and difficult than other web applications. However, considering the many conveniences offered by the system, it has been concluded that this problem can easily be ignored. Django developers are expected to update the architecture to avoid these problems in future projects.

References

[1] Anonymous, 2020, Python (programming language) [online], Wikipedia, Web address: https://tr wikipedia org/wiki/Python_(programlama_language) [Date of Visit: 10th May, 2020]

[2] Malkoç B, 2012, Temel Bilimler ve Mühendislik Eğitiminde Programlama Dili Olarak Python, Akademik Bilişim'12 - XIV Akademik Bilişim Konferansı Bildirileri, Uşak , pp 201-210
[3] Anonymous, 2020, Guido van Rossum [online], Wikipedia, Web address: https://tr wikipedia org/wiki/Guido_van_Rossum [Date of Visit: May 12th, 2020]
[4] Ayvaz U, Çoban A, Gürüler H, Peker M, 2016, Python Dilinin Öznitelikleri, Programlama Eğitiminde ve Yazılım Dünyasındaki Yeri , XVIII AKADEMİK BİLİŞİM KONFERANSI, Aydın, pp 189-194
[5] Anonymous, 2020, TIOBE Index for May 2020 [online], Tiobe, Web address: https://www tiobe com/tiobe-index/ [Date of Visit: May 12th, 2020]
[6] Anonymous, 2020, Why Django, Django, Web address: https://www djangoproject com/start/overview/ [Date of Visit: August 1st, 2020]
[7] Kasapoğlu M, Yazılım Test Planlama Ve Raporlama Aracı'nın Tasarlanması Ve Kodlanması, Master, Hacettepe Üniversitesi Fen Bilimleri Enstitüsü, 2019
[8] Javadi T, Barışçı N, 2018, Duygusal Twitter Analizi, SETSCI ISAS 2018 - Ist International Symposium on Innovative Approaches in Scientific Studies, Antalya, 380-385
[9] Aslan A, Bilgin N 2018, Otomatik Kontrol Eğitimi Desteklenmesi İçin Etkileşimli Web Tabanlı Kontrol Uygulamaları Geliştirilmesi, Otomatik kontrol Türk Milli Komitesi TOK, Kayseri, pp 17-23
[10] Karaş İ, Baz İ, 2007, Üniversite Bölüm Bilgi Sistemi, Akademik Bilişim'07 - IX Akademik Bilişim Konferansı Bildirileri Dumlupınar Üniversitesi, Kütahya, pp 473-480
[11] Ergin İ, Akseki B Lisansüstü Eğitimde Kullanılan Öğrenci Bilgi Sistemi, Eğitim ve Öğretim Araştırmaları Dergisi (JRET) 2012; 1:364-380
[12] Özel H, 2018, Django Nedir [online], medium com, Web address: https://medium com/@halilozel1903/django-nedir-52b29ebb7298 [Date of Visit: 13 Mayıs 2020]
[13] Anonim, 2020, Build fast, responsive sites with Bootstrap [online], Bootstrap, Web address: https://getbootstrap com/ [Date of Visit: 10th May 2020]

Comparison of SISO and I-Metra MIMO Channel Capacities Under Rayleigh Fading Channel

F.Yavuz ILGIN[1]

[1]Erzincan Binali YILDIRIM University, Turkey, fyilgin@erzincan.edu.tr

Abstract - **In today's wireless communication systems, the increase in the size of the data has revealed the importance of the transmission capacity of the communication channels capacity. The use of multiple antennas to the transceiver and receiver has emerged as a solution to the transmission capacity. In this study, it is aimed to compare the channel capacities of MIMO and SISO communication channels. In addition, For MIMO systems, the theoretical analysis of the channel capacitance for different antenna numbers has been made and supported by the simulation studies**

Keywords **- Communication channel, Fading channels, MIMO, OFDM, Rayleigh probability density function.**

I. INTRODUCTION

In recent years, it has been observed that the data sizes transmitted in wireless communication systems have increased gradually[1]. The main reasons for this increase are the use of social networking applications, video conferencing and other wireless communication applications[2]. In order to increase the transmission capacity of communication channels, some multiplexing techniques have been proposed[3]. These methods can send multiple sources from the same channel or apply different coding techniques to the data. In addition, the use of multiple antennas in the receiver and transmitter has become a common method for the efficient use of communication channels[4]. These systems, called Multiple Input Multiple Output (MIMO), increase the channel capacity and the amount of data transmitted by making space diversification[5,6]. Besides MIMO have been produced to reduce the loss of performance of multi-path fading in wireless communications and to meet high information rate demands for future wireless communications with extended spectral efficiency. Much work has been done for MIMO systems, some of which are based on channel capacity. Thus, different studies have been carried out to measure the capacity of the channel. These studies have been applied to different communication channels for OFDM systems[4,7]. Spatial correlation effects for MIMO channel capacity, and studies investigating noise immunity of these channels, are also available in the literature[8,9].

In this study, MIMO and SISO channel capacity were calculated theoretically. Theoretical calculations and simulation results were made in Rayleigh channel environment. In addition to evaluate MIMO channel capacity different antenna combinations were used in the simulation studies.

The remaining of the paper is organized as follows. In Section II, we describe the communication channel types. The SISO channel model and its capacity are described in section III. The MIMO channel model ans its capacity are described in section IV. The simulation results for studies are given in section V. In Finally, Section VI concludes the paper.

Throughout this paper, we use boldface letters for matrices and normal letters for column vector. In addition † means transpose conjugate.

II FADING COMMUNICATION CHANNELS

Generally wireless communication channels can be evaluated in two groups such as large-scale fading and small-scale fading[10]. Large-scale faing is due to path loss and signal attenuation over long distances. Therefore, the variation of such attenuation is not instantaneous. Large-scale fading is affected by evident terrain contours (such as forests, hills, forests, clumps of buildings, etc.) between the radio transmitter. Such amplitude changes in the receiver are often described as "shadowing''. Large-scale fading statistics provide a way to calculate the path loss estimate as a function of distance. This is defined as mean-path loss and average-varying log-normal distribution[11].

Small-scale fading occurs as a result of small changes in the spatial position between a receiver and a transmitter. It indicates random changes in the amplitude and phase of the transmitted signal. Small-scale fading occurs in two groups, ie the time propagation (or signal dissipation) of the signal and the variance behavior of the channel over time. For communication applications, the channel is time-varying because the motion between the transmitter and receiver causes changes in the propagation path. The large scale and small scale fading are shown in Figure 1. Fading in communication systems varies according to the distance between the receiver and the transmitter and the type of the environment. Different channel models are used in the literature. In section 2, presents some of the most important channel models.

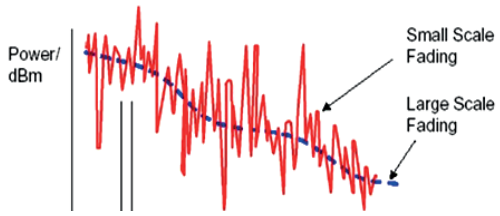


Figure 1. Small scale and large scale fading

A. Rayleigh Fading Channel Model

Rayleigh distribution is often used to model multi-path faing in situations where there is no direct line-of-sight (NLOS), as shown in Fig. 2. The Probability Distribution Function (PDF) for the Rayleigh distribution is given Equ 1[7].

$$p_a(a) = \frac{2a}{\Omega}\, exp\left(-\frac{ma^2}{\Omega}\right) \quad (1)$$

Where $\Omega = a^2$ is the mean-square value of the fading amplitude.

B. Nakagami-n Fading Channel Model

The Nakagami-n distribution is also known as the Rice distribution. It is used to model propagation paths consisting of a strong line of sight (LOS) component and a large number of random weak components. The fading amplitude of the channel is defined by the following PDF[12],

$$p_a(a) = \frac{2(1+n^2)e^{-n^2}}{\Omega}\, exp\left(-\frac{(1+n^2)a^2}{\Omega}\right) I_0\left(2na\sqrt{\frac{1+n^2}{\Omega}}\right) \quad (2)$$

where n is the fading parameter, which can take values in the range of 0 to ∞, and I_0 is modified Bessel function of the first type . The Nakagami-n distribution represents Rayleigh fading at n= 0 and non-fading at n=∞.

C. Nakagami-q Fading Channel Model

Nakagami-q distribution is observed in radio channel due to strong ionospheric clipping. The Nakagami-m distribution is also known as the Hoyt distribution. The PDF of the Hoyt distribution is as follows[13].

$$p_a(a) = \frac{2(1+q^2)a}{q\Omega}\, exp\left(-\frac{(1+q^2)^2a^2}{4q^2\Omega}\right) I_0\left(\sqrt{\frac{(1-q^4)a^2}{4q^2\Omega}}\right) \quad (3)$$

where q is the fading parameter, which can take values between 0 and 1.

D. Nakagami-m fading channel

The PDF of the Nakagami-m distribution is given Equ. 4[14].

$$p_R(r) = \frac{2m^m r^{2m-1}}{\Omega^m \Gamma(m)}\, exp\left(-\frac{mr^2}{\Omega}\right) \quad (4)$$

Where m and Γ are the fading parameter that can take values between 0-∞ and the modified Bessel function respectively. The Nakagami-m distribution represents Rayleigh fading at m = 0 and non-fading at m = ∞. There is an inverse correlation between the fading parameter and the fading intensity. Where Ω denotes the mean-square value of the fading amplitude. Nakagami-m distribution is the best probability distribution for modeling terrestrial and indoor multi-path propagation.

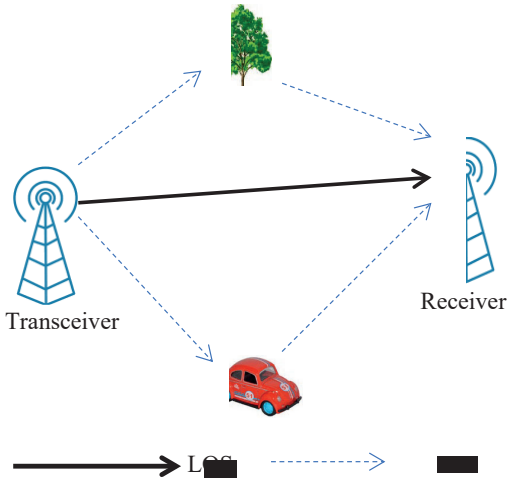


Figure 2. Representation LOS and NLOS

III. SISO CHANNEL MODEL

A communication system consisting of a receiver, transmitter and communication channel is simply modeled as follows.

$$y = Hx + n \quad (5)$$

Where y is received signal. H is channel matrix and n and x are communication signal and additive White Gaussian noise respectively. SISO channel capacity is theoretically expressed as follows.

$$C = E_H[log_2(1 + \rho|h_{11}|^2)] \quad (6)$$

Where ρ represents the signal to noise ratio in dB. If the communication channel model is Rayleigh, $|h_{11}|^2$ is similar to

the Chi-Square distribution with a two-degrees of freedom. Therefore, channel capacity can be expressed as Equ. 7.

$$C = E_H[log_2(1 + \rho\mathcal{X}_2^2)] \quad (7)$$

Where $\mathcal{X}_2^2$ is a chi-square distributed random variable with two degrees of freedom.

IV. MIMO CHANNEL MODEL AND ITS CAPACITY

MIMO communication systems are widely used in order to meet the increasing capacity and transmission speed requirements for wireless communication systems. MIMO systems are a modulation and multiplexing method that can use spatial dimension effectively in addition to frequency and time dimensions. Therefore, it uses the spectrum much more efficiently than traditional modulation and communication techniques[4]. In MIMO systems, the serial data to be transmitted is transmitted in parallel by dividing it into multiple sub-carriers. MIMO communication systems are modeled and designed with the assumption that the channels of subcarriers are theoretically uncorrelated. Therefore, the channel capacity increases depending on the antenna nearness[15]. However, there are many factors that can change the channel capacity in physical propagation environments. At the very beginning, the spatial correlation between the antennas can be shown. To avoid this, the distance between the antennas must be sufficiently far, but this is not easy in practice. The MIMO system is shown in Figure 3.

Where Tx and Rx are show transmit and receiver antennas respectively. Where, according to the Shannon channel capacity theory, the capacity can be increased without increasing the bandwidth. The most important part for understanding and handling MIMO channel capacity is the channel matrix.

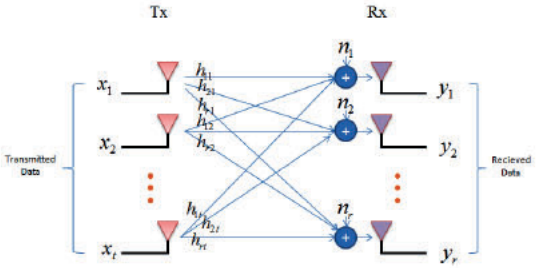


Figure 3. MIMO channel model

The channel matrix for MIMO channel is given below.

$$\boldsymbol{H}(\tau,t) = \begin{bmatrix} h_{1,1}(\tau,t) & \cdots & h_{1,n_t}(\tau,t) \\ \vdots & \ddots & \vdots \\ h_{n_r,1}(\tau,t) & \cdots & h_{n_r,n_t}(\tau,t) \end{bmatrix} \quad (8)$$

Where $h_{1,1}(\tau,t)$ defines the channel coefficient from the 1st transmitting antenna to the second receiving antenna. In this case MIMO channel capacity is expressed as in Equ. 9.

$$C = log_2\left(det\left[I_{n_r} + (p/n_t)\right]HH^{\dagger}\right)\frac{b}{s}Hz \quad (9)$$

Where I_{n_r} means $n_r x\, n_t$ identity matrix.

$$C > \sum_{k=1}^{n} log_2\left([1 + (p/n)]\mathcal{X}_{2k}^2\right)\frac{b}{s}Hz \quad (10)$$

The channel capacity for the $n_r x\, n_t$ case in terms of the independent $\mathcal{X}_{2k}^2$ variable with two-degrees of freedom is as follows.

$$C = log_2(det[1 + (\rho\mathcal{X}_{2k}^2)])\frac{b}{s}Hz \quad (11)$$

V. NUMERICAL RESULTS

In the simulation studies, Monte Carlo simulation was performed for MIMO channel capacities, for different antenna numbers. The horizontal axis indicates SNR(dB) and the vertical axis indicates the channel capacity(b/s/Hz). In the graphs, nTx and nRx indicate the number of transmitting antennas and the number of receiving antennas respectively. In Fig. 4 the antenna numbers were selected mixed. Where nTx =1, nRx =1 also refers to the SISO system. Looking at the graphs, there is a direct correlation between noise level and channel capacity, which can be explained by the Shannon channel capacity theorem. Rayleigh channel model was used in the simulations.

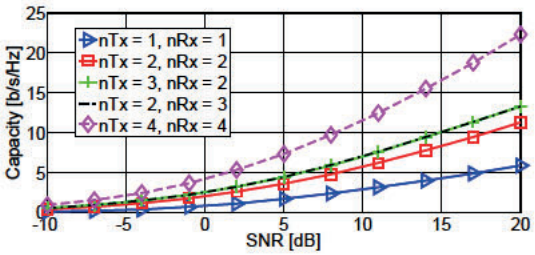


Figure 4. Capacity of a MIMO system for different num of Tx-Rx antennas

In Fig. 5, the number of transmitting antennas is 2. Thus the channel capacity according to the number of different antennas in the receiver is shown.

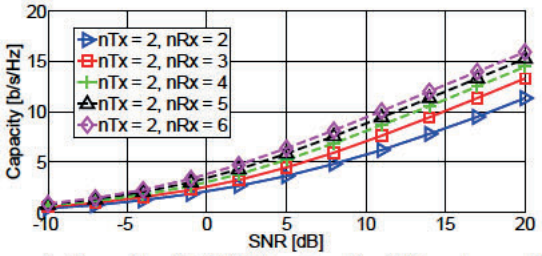


Figure 5. Capacity of a MIMO system for different num of Tx-Rx antennas

The number of antennas in the transmitter 4 is given in figure 6. Thus the channel capacity according to the number of different antennas in the receiver is shown. In addition, the

condition that contains the most antennas in the transmitter and receiver is given in Figure 6. The channel capacity for a 4 x 6 MIMO system (nTx = 4, nRx = 6) appears to be approximately 28 b/s/Hz.

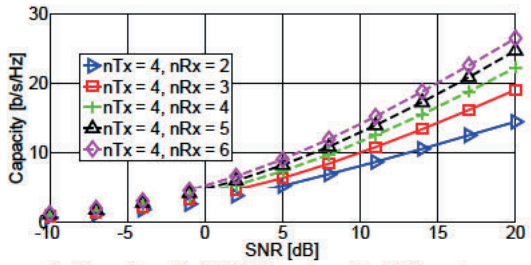


Figure 6. Capacity of a MIMO system for different num of Tx-Rx antennas

According to the simulation studies, channel capacities are highly dependent on noise. The 4x2 MIMO system offers 20 channels capacity at 20 b/s/Hz dB noise level. But in practice, the noise level is not so low. Besides, due to the noise uncertainty factor, the channel capacity may be slightly lower.

VI. Conclusion

In this study, channel capacity of MIMO communication systems is examined. Channel capacity analysis for SISO and MIMO communication systems is theoretically performed. As can be seen from the numerical results section of the study, the channel capacity for MIMO communication systems varies with the number of antennas. But the antennas must be uncorrelated. To ensure this condition the distance between the antennas must be long enough.

References

[1] Çiflikli C, Ilgin FY Covariance Based Spectrum Sensing with Studentized Extreme Eigenvalue *Tech Gaz* 2018;25:100–6

[2] Wang J, Ghosh M, Challapali K Emerging cognitive radio applications: A survey *IEEE Commun Mag* 2011;49:74–81 doi:10 1109/MCOM 2011 5723803

[3] Souid I, Ben Chikha H, Dayoub I, Attia R MIMO relaying networks for cooperative spectrum sensing: False alarm and detection probabilities *Phys Commun* 2017;25:194–200 doi:10 1016/j phycom 2017 07 006

[4] Zheng L, Tse DNC Diversity and multiplexing: A fundamental tradeoff in multiple-antenna channels *IEEE Trans Inf Theory* 2003;49:1073–96 doi:10 1109/TIT 2003 810646

[5] Lu L, Li GY, Swindlehurst AL, Ashikhmin A, Zhang R An overview of massive MIMO: Benefits and challenges *IEEE J Sel Top Signal Process* 2014;8:742–58 doi:10 1109/JSTSP 2014 2317671

[6] Dahlman E, Parkvall S, Skold J 4G: LTE/LTE-Advanced for Mobile Broadband 2013 doi:10 1016/C2013-0-06829-6

[7] Kang M, Alouini MS Capacity of MIMO Rician channels IEEE *Trans Wirel Commun* 2006;5:112–22 doi:10 1109/TWC 2006 1576535

[8] Sheikh TA, Bora J, Hussain MA Capacity maximizing in massive MIMO with linear precoding for SSF and LSF channel with perfect CSI *Digit Commun Networks* 2019 doi:10 1016/J DCAN 2019 08 002

[9] Griffiths HD, Baker CJ, Sammartino PF, Rangaswamy M MIMO as a Distributed Radar System *MIMO Radar Signal Process.*, 2008, s 319–63 doi:10 1002/9780470391488 ch8

[10] Yacoub MD, Bautista JEV, Guerra de Rezende Guedes L On higher order statistics of the Nakagami-m distribution IEEE *Trans Veh Technol* 1999;48:790–4 doi:10 1109/25 764995

[11] Liang W, Le KN A review of selection combining receivers over Nakagami-m correlated fading channels Digit Signal Process 2018;83:180–201 doi:10 1016/J DSP 2018 08 018

[12] Alouini MS, Goldsmith AJ Adaptive modulation over Nakagami fading channels *Wirel Pers Commun* 2000;13:119–43 doi:10 1023/A:1008979107539

[13] Bandari SK, Vakamulla VM, Drosopoulos A GFDM/OQAM performance analysis under Nakagami fading channels *Phys Commun* 2018;26:162–9 doi:10 1016/J PHYCOM 2017 12 008

[14] Beaulieu NC, Cheng C Efficient nakagami-$\mathcal{M}$ fading channel simulation *IEEE Trans Veh Technol* 2005;54:413–24 doi:10 1109/TVT 2004 841555

[15] Shieh W, Bao H, Tang Y Coherent optical OFDM: theory and design *Opt Express* 2008;16:841–59 doi:10 1364/OE 16 000841

Capture Date Estimation of Historical Images By Using Deep Learning

Y.CIMTAY[1] and H.G. İLK[2]

[1] HAVELSAN, Image and Signal Processing Group, Ankara/Turkey, ycimtay@havelsan.com.tr
[2] Ankara University, Electrical and Electronics Engineering, Ankara/Turkey, ilk@ieee.org

***Abstract* - The issue of automatic estimation of the capture date of historical photographs have seldom been discussed in the past. This paper implements a deep learning-based automatic prediction method. For the training of deep learning model, we created an historical image dataset which consists of photographs which were captured from different cities of Turkey. This dataset 's date interval is from the 1880s to the 2010s. Instead of extracting specific feature sets like traditional methods do, we employ a CNN model which can extract hidden and distinguishing features from raw images. Results show that our proposed model is superior to human visual inspection.**

***Keywords* – Convolutional Neural Network, Culture, Fashion, Urbanization, Historical Photos.**

I. Introduction

It is possible to see large datasets in the literature such as YFCC100M [1] and ImageNet [2] which are widely used for computer vision applications. This kind of comprehensive and qualified datasets are crucial especially for deep learning studies. However, in terms of automatic capturing date estimation, there is not enough data especially there is lack of data belong before 2000s. Besides, capturing date estimation is a challenging and interesting task for historians and states to use in historical research and some detective issues. Existing approaches deal with special objects like persons [3,4], or historical documents [5,6]. In addition, some extends the dataset and try to develop methods completely based on huge and rich raw data [7]. In [7] authors believe in that huge datasets which cover all kinds of concepts is necessary for the training of convolutional neural networks. In this study, we collected lots of historical images from different regions of Turkey. We rely on the dressing style of people, vehicle types and architectural style of buildings except the historical ones.

II. Related Work

The first study related capturing date estimation of historical images is conducted in [8]. It is based on overlapping many temporal images of the same region for sorting the historical image collection. The study in [9] focuses on patch groups of street and car scenes in order to model the style differences across time. The studies conducted in [5,6] dealt with estimation date of historical documents. [5] emphasized the importance of contour and strokes in the documents. Besides, [6] combined convolutional neural network and optical character recognition. The study in [7] develops a fully independent deep learning-based system which is based on using a huge dataset which composed of raw historical images. Like [7], the study in [10] combines various color descriptors to model the historical color film processes. Like our study, the study in [4] authors model the diversity of human appearance and dressing style to estimate the date of photographs.

III. Dataset

In this study we created an historical dataset by using the historical images which can be found in google. It covers the date from 1880s to 2010s. The images are from different regions/cities of Turkey and totally we have 885 images. We have segmented the human and vehicle regions and increase the number of images to 3176. Number of images belong to each 10 years period are given in Table 1.

Table 1: Number of images belong to each 10-years period.

10-Years period	Number of images
1880s	168
1890s	164
1900s	208
1910s	168
1920s	311
1930s	291
1940s	322
1950s	221
1960s	297
1970s	294
1980s	176
1990s	224
2000s	152
2010s	180

The regions which the images are captured from are but not limited to Istanbul, Bursa, Elazığ, Afyon, Adana, Niğde, Konya, Trabzon, Tokat, İzmir and Mardin. The procedure of photo selection is,

- Blurred photos should not be selected
- Black-and-White photos can be selected
- Photographs that contain intense noise that will prevent the details from being seen should not be selected.
- The label indicating the date of the photograph should be compared with the real historical development, photographs that are thought to be given a very obviously incorrect date should not be selected.
- It should be shown that the photographs taken include people or transport vehicles (phaeton, horse carriage, automobile, van, etc.).

- Attention should be paid to the characteristics of the building in photographs that do not contain people or vehicles. Photographs should be selected in which the buildings are in an architecture suitable for the historical structure. For example, if there is a historical building in a photograph (eg. Sumela Monastery), this photograph should be taken into a separate group. For example, photos with apartment buildings increasing after the 70's are suitable to be selected. Or, photos with mansion style buildings from older years are suitable for selection.
- Photographs should generally be chosen from city centers or regions whose development level is not too low. For example, photographs that include the dress style of the 1950s in the 1980s should not be selected.
- Photographs should be grouped according to 10-year intervals. For example, 1900-1910, 1980-1990.

There are totally 37 different cities of Turkey in the dataset. Three example images are shown in Figure 1 where "a" is from 1880s, "b" is from 1950s and "c" is from 1980s and "d" is from 2010s. There are obvious differences in the clothing styles of humans and the style of cars. Therefore, our photo selection procedure is very important in terms of selecting the target imagery, high quality features and so increasing the classification accuracy.

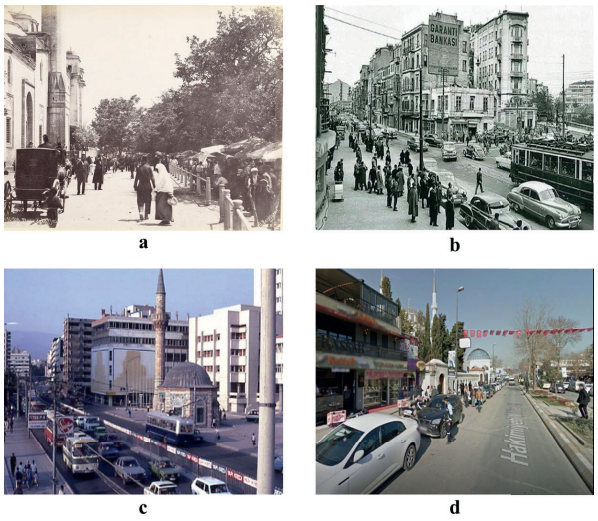


Figure 1: Examples photos.

IV. Proposed Method

After creation of the dataset, we resized the images to be 256x256 in order to bring a standard for input of proposed network. We propose a CNN (Convolutional Neural Network) model. CNN is a state-of-the-art network model which is superior to traditional methods in terms of data classification, especially when raw data is used [11]. We applied data augmentation in order to increase the number of input samples. After several tries the optimal final network structure, we use is given in Table 2. It shows the network layers and the output shape of the data after each layer. We use a Xception architecture as the base model. The model plot is given in Figure 2.

Table 2: Layer description of CNN.

	Output Shape
	(None, 256, 256, 3)
	(None, 8, 8, 2048)
	(None, 2048)
dropout	(None, 2048)
dense	(None, 5)

The image dataset was split into training and test sets. The training parameters of proposed CNN is given in Table 3.

Table 3: Training parameters of CNN.

Parameter	**Value**
optimizer	adam
loss	categorical_crossentropy
shuffle	True
Number of epochs	300
batch_size	64

In order to increase the separability of classes, instead of using 10-years periods we regrouped the images based on revolutions and important events happened in Turkey such as the foundation of Turkish Republic. Therefore, we had 1880-1920, 1920-1950, 1950-1980, 1980-2000, 2000-2010 date intervals as 5 classes. We split each group into training and test sets. We employed a random selection method and split with ratio of 0.25 and 0.75 to training and test sets respectively.

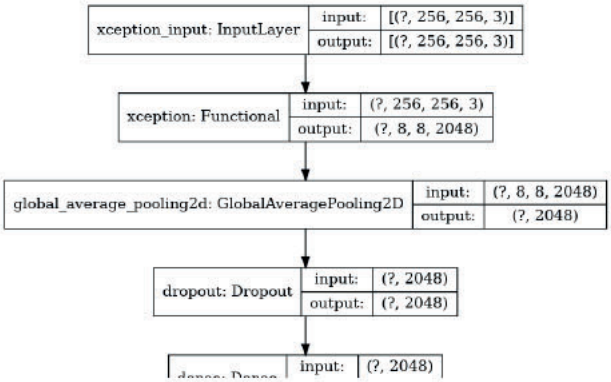


Figure 2: Model plot.

V. RESULTS

The classification results are given as a confusion matrix in Table 4. Overall accuracy is calculated as 68.01%.

Table 4: Confusion matrix.

Target Class							
	1880-1920	105	52	19	1	0	0.59
	1920-1950	18	191	14	1	7	0.82
	1950-1980	13	38	138	6	8	0.67
	1980-2000	11	14	20	45	10	0.45
	2000-2010	2	10	7	3	61	0.73
		0.70	0.62	0.69	0.78	0.70	0.68
		1880-1920	1920-1950	1950-1980	1980-2000	2000-2010	

The distributions of classified samples are generally very logical. For instance, the number of (1880-1920) samples which are classified as (1920-1950) is high compared to the newer date intervals. Similarly, the number of (1950-1980) which are classified as (1920-1950) is higher compared to the (2000-2010) date interval.

The column on the far right shows the recall scores for each class, and the gray raw on the bottom shows the precision scores for each class. The best recall score is achieved for 1920-1950 and the worst is for 1980-2000 interval. The best precision score is achieved for 1980-2000 and the worst is for 1920-1950 interval.

We also calculated the MSE (mean square error) for estimated date. Table 5 shows the benchmark MSE scores for labelling by human (an historian) and classification result of proposed method. It can be seen from the table that for all date intervals except 1980-2000, our proposed model is superior to human evaluation. This is a promising result and shows that the proposed approach can be improved.

Table 5: Benchmark MSE scores, human & model.

Date Interval	Proposed Model	Human evaluation
1880-1920	7.33	9.04
1920-1950	4.71	5.45
1950-1980	8.91	9.76
1980-2000	7.18	5.11
2000-2010	3.37	3.43

VI. CONCLUSION

In this study, we have implemented a deep learning model for capture date estimation of historical photographs. The dressing style and the vehicles are the most changing issues for human and daily life. However, this differentiation is highly depending on the culture and lifestyles of society. Therefore, we prepared a dataset which is composed of the photographs of different regions of Turkey. This study is a base study for capturing date estimation of photographs for Turkey. We created a base dataset to be used further studies.

We achieved promising precision and accuracy classification results. We believe that increasing the number of samples in dataset will increase the accuracy of each classes so the overall accuracy. In addition, we can see how artificial intelligence is superior to human evaluation. This study is very important in terms of creating a base for later work about capturing date estimation of historical photographs.

Increasing the samples in dataset and using deeper networks is going to most probably increase the accuracy.

ACKNOWLEDGMENT

The authors would like to thank Polat BEDEL (who we lost last year) for his efforts in data collection and dataset creation and History Teacher Saliha CIMTAY for labelling the imagery by visual inspection.

REFERENCES

[1] B Thomee, D A Shamma, G Friedland, B Elizalde, Ni,K Poland, D ,Borth,D , Li, L J : "YFCC100M: the new data in multimedia research " *Commun. ACM* 59(2), 64–73 (2016)

[2] O Russakovsky, J Deng, H Su, J Krause, S Satheesh, S Ma, Z Huang, A Karpathy, et al : "Imagenet large scale visual recognition challenge" *Int. J. Comput. Vis.* 115(3), 211–252 (2015)

[3] S Ginosar, K Rakelly, S Sachs, B Yin, A Efros, "A century of portraits: a visual historical record of American high school yearbooks," *In: IEEE International Conference on Computer Vision Workshops*, pp 1–7 (2015)

[4] Salem, T , Workman, S , Zhai, M , Jacobs, N : Analyzing human appearance as a cue for dating images In: 2016 IEEE Winter Conference on Applications of Computer Vision, pp 1–8 (2016)

[5] He, S , Samara, P , Burgers, J , Schomaker, L : Image-based historical manuscript dating using contour and stroke fragments Pattern Recogn 58, 159–171 (2016)

[6] Y Li, D Genzel, Y Fujii, A C Popat, “Publication date estimation for printed historical documents using convolutional neural networks,” *In: Third International Workshop on Historical Document Imaging and Processing*, pp 99–106 (2015)
[7] E Müller, M Springstein, R Ewerth, “When Was This Picture Taken?” – Image Date Estimation in the Wild *In: Jose J. et al. (eds) Advances in Information Retrieval. ECIR 2017. Lecture Notes in Computer Science*, vol 10193 Springer, 2017
[8] G Schindler, F Dellaert, S B Kang, “Inferring temporal order of images from 3D structure,” *In: IEEE Conference on Computer Vision and Pattern Recognition*, pp 1–7 (2007)
[9] Y Jae Lee, A A Efros, M Hebert, “Style-aware mid-level representation for discovering visual connections in space and time,” *In: IEEE International Conference on Computer Vision*, pp 1857–1864 (2013)
[10] F Palermo, J Hays, A A Efros, “Dating historical color images,” *ECCV 2012. LNCS,* vol 7577, pp 499–512 Springer, Heidelberg (2012)
[11] Y Cimtay, E Ekmekcioglu and S Caglar-Ozhan, "Cross-Subject Multimodal Emotion Recognition Based on Hybrid Fusion*," in IEEE Access*, vol 8, pp 168865-168878, 2020

Performance Evaluation of Average and Median Filters for Gaussian Noise in Ceramic Tile Images

M. A. GUNGOR

Hitit University, Çorum/Turkey, alparslangungor@hitit.edu.tr

***Abstract*— Today, the ceramic tile industry is a very important sector. After the ceramic tiles are produced, quality control is carried out to prevent defective products from reaching the customer. One of the methods used for quality control is the image processing application. Noise reduction algorithms in the image processing application are used to reduce noise in the image. In this paper, we analyze the effects of the average and median filters on the ceramic tile images with Gaussian noise. First, noisy images with different noise levels are obtained. Then, the filters with different properties are applied to noisy images. The filtered images are analyzed using the quality metric. The results show that the median filter performs better performance than the average filter for both low and high noisy images.**

***Keywords*— Image Denoising, Gaussian Noise, Median Filter, Average Filter, Ceramic Tile**

I. INTRODUCTION

When quality control of ceramic tiles is done with image processing techniques, the computers are used so that human intervention is minimized or not [1]. The target is to find defects such as cracks, edges, holes, and pinholes [2, 3]. Many studies have been done in the literature to detect the defects found in ceramic tiles [1-13]. The goal is to identify defective products and classify them if necessary.

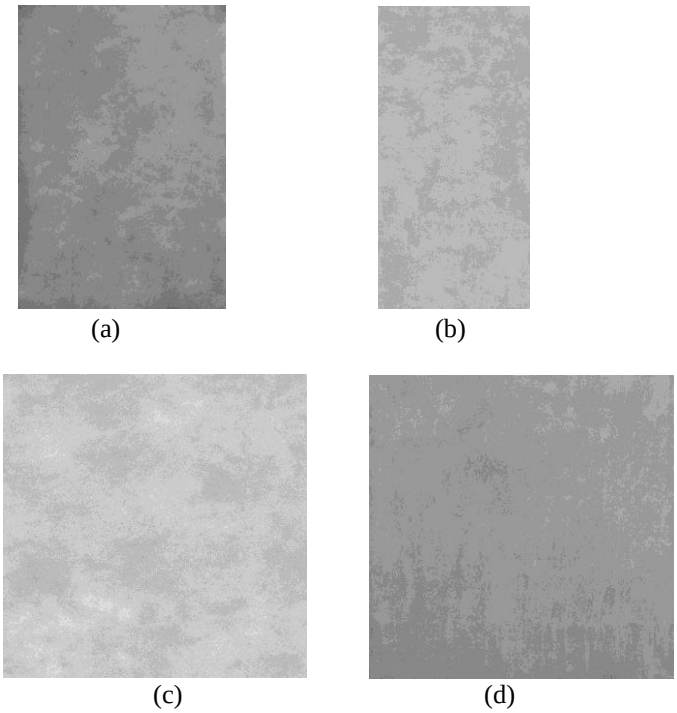
For processing the image, first, the captured ceramic tile image in the RGB format is converted to a gray-scale format. Then, various operations are performed for more effective quality control. In the image processing application, one of the most important operations is the noise reduction operation. The better the noise reduction performance, the better quality control.

There is a wide variety of noises in the images. Gaussian noise arising in an image by factors such as sensor noise, electronic circuit noise can be modeled as a very common noise. [14, 15]. The conventional filters such as the average and median filters are efficient in reducing the noise [14]. The median and average filters have been used for many noise reduction applications of ceramic tile images [1, 4, 6].

In this paper, we analyze the effects of average and median filters on the ceramic tile images with Gaussian noise. We use two types of average filters with different coefficients and five types of median filters with different window sizes. Using these filters, we filter twelve ceramic tile images with different noise levels. We use peak signal-to-noise ratio (PSNR) as an image quality metric. Thus, we compare the filters according to their noise reduction performance.

II. MATERIALS AND METHODS

In this paper, we employ four types of reference ceramic tile images, shown in Fig. 1.



(a) (b) (c) (d)

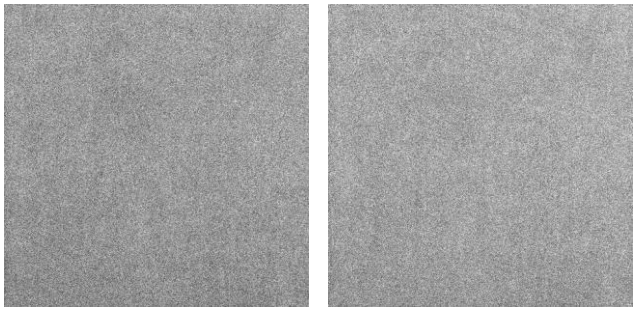
Figure 1: Reference images (a) image 1 (b) image 2 (c) image 3 (d) image 4

The images shown in Fig. 1 are the gray-scale format and obtained from the images in the RGB format, based on Eq. (1), using in the image processing toolbox of MATLAB [16].

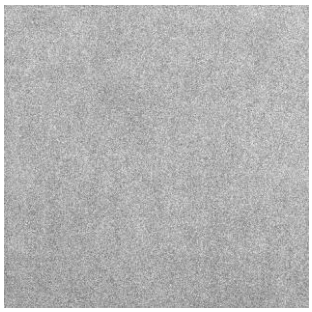
$$rgb2gray=0.2989R + 0.5870G + 0.1140B \quad (1)$$

In this paper, we analyze homogeneous images. The images shown in Fig. 1 do not contain edge information. These images are reference images. To obtain noisy images, noise is added to the reference image. In this study, noisy images are obtained by adding Gaussian noise. Three different noise levels, called noise level 1, noise level 2, and noise level 3 are used. Noise level 1 indicates the least noise, while noise level 3 indicates the highest noise. For example, Fig. 2 shows the noisy images of image 4 in Fig. 1. Comparing Fig. 2 with Fig1d, the noisy

image 4 with noise level 3 is the most distorted image among the images given in Fig. 2.



(a) (b)



(c)

Figure 2: Noisy images of image 4 for (a) noise level 1 (b) noise level 2 (c) noise level 3

The purpose of the denoising filters is to reduce the noise in the noisy image and obtain the image most similar to the reference image. We use two types of denoising filters, median and average. The median filter uses the middle value in the window. First, the pixel values in the window are sorted numerically and then the median value is obtained. Finally, the central pixel value in the window is replaced with the obtained median value. An example of median filtering using a 3*3 window size is shown in Fig. 3. In Fig. 3, the central pixel value of 20 is replaced with the median value 50.

22	10	78	19	45
63	10	50	90	88
30	80	20	40	96
96	30	70	60	11
78	44	63	15	56

Figure 3: Calculating the central pixel value using the median filter with a 3*3 window size

The average filters used in this paper are shown in Fig. 4. Fig. 4a shows the standard average filter, while Fig. 4b shows the weighted average filter. The standard average filter is used to average the pixels under the window. All coefficients of the filter are 1 and the entire image is divided by 9. The weighted average filter is used to indicate that pixels are multiplied by different coefficients, so some pixels are given more importance than others. In the window shown in Fig. 4b, the pixel at the center of the window is multiplied by a higher value than any other, thus giving this pixel more importance in the calculation of the average.

$(1/9)*$

1	1	1
1	1	1
1	1	1

(a)

$(1/16)*$

1	2	1
2	4	2
1	2	1

(b)

Figure 4: The average filters (a) standard average filter (b) weighted average filters

After filtering the noisy images, we use PSNR as an image quality metric. The PSNR is most easily defined via the mean squared error (MSE) using Eq. (2)

$$MSE = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}\left[I_r(i,j) - I_f(i,j)\right]^2$$

$$PSNR = 10\,log_{10}\left(\frac{MAX_I^2}{MSE}\right) \quad (2)$$

where m and n are the size of the image, MAX_I is the maximum possible pixel value of the image, I_r and I_f are the reference and filtered images, respectively. The higher PSNR value indicates a better quality image after filtering.

III. Results and Discussion

In this paper, four types of reference images are employed. Gaussian noise is added to the reference images to obtain noisy images. Three different noise levels are used; thus twelve noisy images are obtained. The obtained twelve noisy images are filtered with two types of average filters shown in Fig. 4 and the median filters whose sizes are 3*3, 5*5, 7*7, 9*9, and 11*11. Finally, eighty-four filtered and twelve noisy images are analyzed using the PSNR metric. Fig. 5 shows the PSNR values versus noise levels for all images.

As shown in Fig. 5, all filters improve the noisy images. Of all filters, the median filter with a 3*3 window size gives the worst results. Comparing the median filters, the larger the window size, the better quality images are obtained for the images with low noise levels. As the noise level increases, median filters of different window sizes have almost the same results. Comparing the average filters, the standard average filter is either better than the weighted average filter or is almost the same as the weighted average filter.

When the median and average filters are compared, the median filter has better performance than the average filter for the low noise levels. For the high noise levels, the quality of the images obtained from both kinds of filters is very close to each other. As a result, the median filter is a good option compared to the average filter for both low and high noisy images.

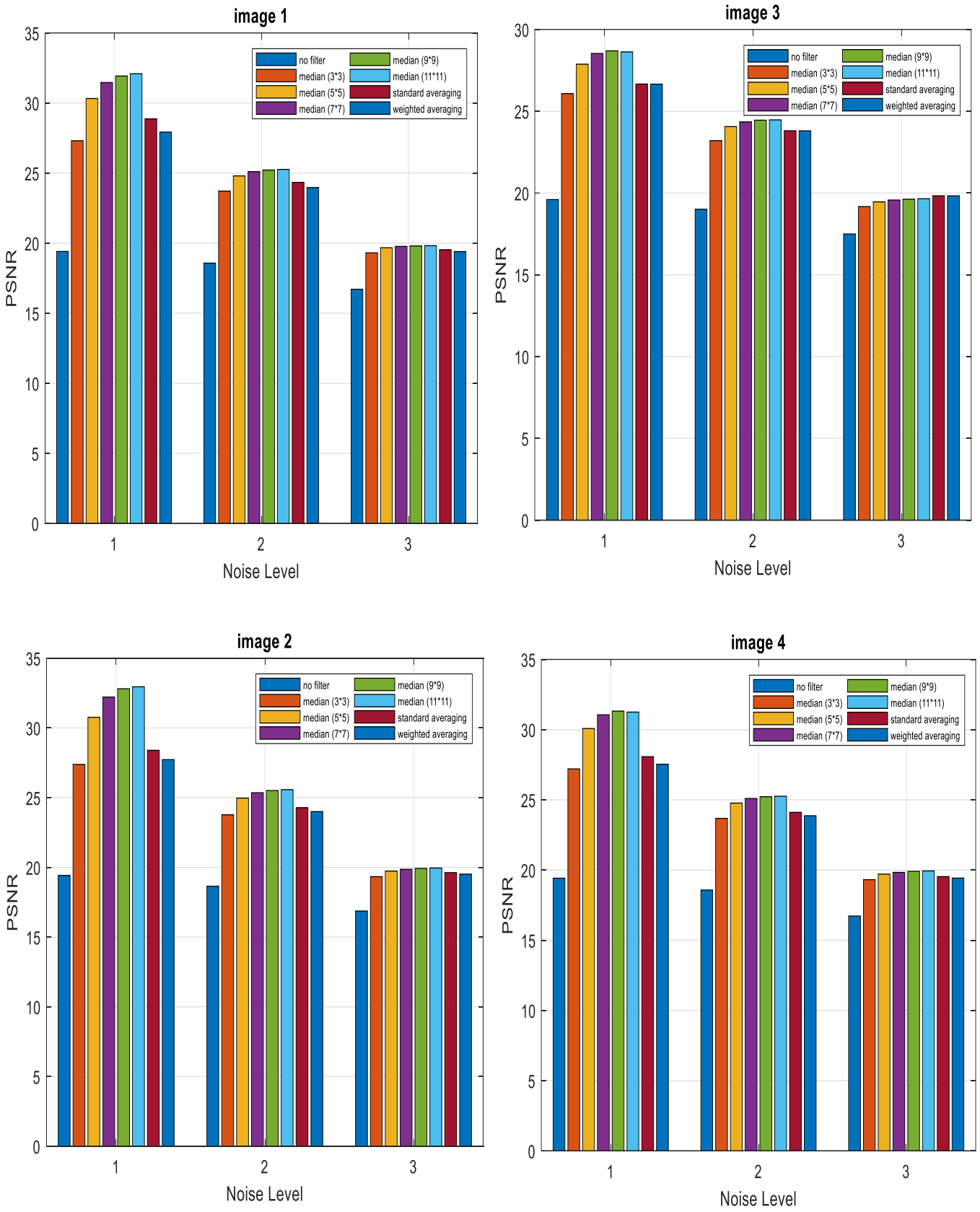


Figure 5: PSNR values versus noise levels

The larger window size causes more processing time, so the smallest possible window size should be selected for the median filter. When Fig. 5 is analyzed, the median filters with 9*9 and 11*11 window sizes give the best performance in the median filters. Thus, the median filter with a 9*9 window size can be selected for denoising. Furthermore, the window size can be further reduced for high noisy images, especially for applications where processing time is critical.

IV. CONCLUSIONS

In this paper, we analyze the average and median filters to remove the Gaussian noise in ceramic tile images. The noisy images with different noise levels are filtered with filters of different sizes and characteristics. The results show the median filter has better performance than the average filter for the reduction of the Gaussian noise in the ceramic tile. Besides, for applications where processing time is very important, if the noise level in the image is high, the median filter with a smaller window size can be preferred.

REFERENCES

[1] Rahaman, G. M., & Hossain, M. (2009). Automatic defect detection and classification technique from image: a special case using ceramic tiles. International Journal of Computer Science and Information Security. 1(1), 22-30.

[2] Karimi, M. H., & Asemani, D. (2014). Surface defect detection in tiling Industries using digital image processing methods: Analysis and evaluation. ISA transactions, 53(3), 834-844.

[3] Czimmermann, T., Ciuti, G., Milazzo, M., Chiurazzi, M., Roccella, S., Oddo, C. M., & Dario, P. (2020). Visual-Based Defect Detection and Classification Approaches for Industrial Applications—A SURVEY. Sensors, 20(5), 1459.

[4] Hanzaei, S. H., Afshar, A., & Barazandeh, F. (2017). Automatic detection and classification of the ceramic tiles' surface defects. Pattern Recognition, 66, 174-189.

[5] Mohan, A., & Poobal, S. (2018). Crack detection using image processing: A critical review and analysis. Alexandria Engineering Journal, 57(2), 787-798.

[6] Shire, A. N., Khanapurkar, M. M., & Mundewadikar, R. S. (2011, November). Plain ceramic tiles surface defect detection using image processing. In 2011 Fourth International Conference on Emerging Trends in Engineering & Technology (pp. 215-220). IEEE.

[7] Ng, H. F. (2006). Automatic thresholding for defect detection. Pattern recognition letters, 27(14), 1644-1649.

[8] Hocenski, Z., Keser, T., & Baumgartner, A. (2007, June). A simple and efficient method for ceramic tile surface defects detection. In 2007 IEEE International Symposium on Industrial Electronics (pp. 1606-1611). IEEE.

[9] Sioma, A. (2020). Automated Control of Surface Defects on Ceramic Tiles Using 3D Image Analysis. Materials, 13(5), 1250.

[10] Latif-Amet, A., Ertüzün, A., & Erçil, A. (2000). An efficient method for texture defect detection: sub-band domain co-occurrence matrices. Image and Vision Computing, 18(6-7), 543-553.

[11] Vasilic, S., & Hocenski, Z. (2006, July). The edge detecting methods in ceramic tiles defects detection. In 2006 IEEE International Symposium on Industrial Electronics (pp. 469-472). IEEE.

[12] Hocenski, Z., & Keser, T. (2007, June). Failure detection and isolation in ceramic tile edges based on contour descriptor analysis. In 2007 Mediterranean Conference on Control & Automation (pp. 1-6). IEEE.

[13] Samarawickrama, Y. C., & Wickramasinghe, C. D. (2017, January). Matlab based automated surface defect detection system for ceremic tiles using image processing. In 2017 6th National Conference on Technology and Management (NCTM) (pp. 34-39). IEEE.

[14] Barbu, T. (2013, December). Variational image denoising approach with diffusion porous media flow. In Abstract and Applied Analysis (Vol. 2013). Hindawi.

[15] Thakur, N., & Devi, S. (2011). A new method for color image quality assessment. International Journal of Computer Applications, 15(2), 10-17.

[16] Matlab, Image Processing Toolbox

.

An Eye Shaped Patch Antenna Operable at Sub-6 GHz Frequency Bands for Fifth Generation Wireless Communication Systems

T.OKAN[1]

[1] University of Turkish Aeronautical Association, Department of Electrical and Electronics Engineering
Ankara/Turkey, tokan@thk.edu.tr

***Abstract* - In this paper, an antenna is designed that is aimed to be able to operate at all three sub-6 GHz frequency bands (2.5 GHz band, 3.5 GHz band and 3.7-4.2 GHz band) specifically allocated for fifth generation (5G) communication. The proposed patch antenna has an eye shaped radiator and it is fed by a microstrip transmission line. The size of the radiator element, the width of the transmission line and the length of the ground plane are determined precisely in order to cover the abovementioned three frequency bands. A substrate material with a relative permittivity of 2.2 is preferred with a copper thickness of 0.035 mm. The presented antenna has an operating bandwidth ($S_{11} \leq$ -10 dB) from 2.195 GHz through 5.095 GHz. It corresponds to a fractional bandwidth of 80%. Besides its wideband property, the antenna has a radiation efficiency that varies between 87% to 96% throughout the operating frequency band. Furthermore, the proposed antenna has a size of 50 mm × 43 mm and a maximum realized gain of 3.2 dBi. It has an omni-directional radiation pattern at the resonance frequency of 4.2 GHz. Finally, the multi-input multi-output (MIMO) design (4×4) of the antenna is also analyzed and great isolation is observed between the two elements which are positioned perpendicular to each other. As a result of all these analyzes, one can say that this antenna is convenient to be used for sub-6GHz bands of 5G applications.**

***Keywords* - Fifth generation (5G), sub-6 GHz, multi-input multi-output (MIMO) antenna.**

I. INTRODUCTION

THE necessity in high data rate and wide bandwidth wireless communication systems have arisen with the increase in the number of mobile users and in the size of shared data among these users. Fifth generation wireless networks are developed to fulfill these demands. However, there is still an uncertainty in many aspects of this technology. For example, the operating frequency haven't determined yet by the authorized institutions in many countries.

5G frequency bands are mainly classified into two as sub-6 GHz bands and millimeter (mm) wave frequency bands. Although an increase in operating frequency brings the benefits of wider bandwidth and higher data rate, it has a disadvantage of shorter-range usability due to higher attenuation. The mm wave frequency bands are 28 GHz, 38 GHz, 45 GHz, 60 GHz and 70 GHz [1]; whereas the sub-6 GHz frequency bands are listed as 2.5 GHz, 3.5 GHz and 3.7-4.2 GHz. Different antenna designs are reported in [1-11] for different operating frequencies. Capability of multi band is an important feature in antenna designs. It prevents using a second antenna for another frequency. Therefore, a multi band antenna decreases the operating costs as well.

Microstrip patch antennas are the most preferred antenna types thanks to their ease in fabrication, compact size and low cost. In [4], a super wideband antenna was designed that operates in 2.59 GHz to 31.14 GHz range. In addition, the radar cross section calculation was also performed for the presented antenna. A quad band 5G antenna was reported in [9], where the presented antenna operated at 28, 38, 45 and 60 GHz frequencies. Frequency selective surface technique was applied to an end fire antenna that operated at 28 GHz 5G band in [8]. A dual band (28 GHz and 45 GHz) elliptical slot patch antenna was presented in [1]. The presented antenna had an efficiency of 85% and 95%, respectively. Another dual band antenna was proposed in [10], the proposed antenna in this study was a dual-feed antenna and it was designed to operate at 38.5 and 73.5 GHz frequencies. The gain and fractional bandwidth values at the abovementioned frequencies were 2.9 and 3 dBi; and 13% and 14%, respectively. In [7], an antenna was designed to operate at sub-6 GHz bands. That antenna was also capable to perform a tri-band band notch character after a few minor changes in the design. Other 5G antennas operating at sub-6 GHz bands were analyzed in [3] and [2]. In [3], a wideband MIMO antenna design was presented; whereas a low-profile dual band antenna was reported in [2]. A flexible coplanar waveguide fed transparent antenna was proposed in [5]. It had a 76% transparency and 42.2% fractional bandwidth percentage. Although all these abovementioned studies are related to 5G antenna applications, some of them only cover the mm wave frequency bands (28, 38, 45, 60 GHz), some have a complex design, some have low efficiency value and some do not cover the all three sub-6 GHz bands.

In this study, a low profile, highly efficient, wideband microstrip antenna is designed that covers all the 2.5 GHz, 3.5 GHz and 3.7-4.2 GHz frequency bands separated for sub-6 GHz 5G applications. The proposed antenna has maximum gain and efficiency values of 3.2 dBi and 92%, respectively. It has an operating bandwidth from 2.2 GHz to 5 GHz which corresponds to 23% fractional bandwidth percentage.

Furthermore, the 2×2 MIMO condition of the antenna is analyzed as well and great performance results are observed from the simulations.

II. Design and Parametric Study of The Antenna

The proposed antenna design has an eye shaped radiator and fed by a microstrip transmission line. The sizes of the elliptical ring radiator and the circle inside it are determined specifically to resonate at the desired frequency. Moreover, the width and length values of the transmission line are also assigned to set the impedance matching of 50 Ohms [12]. The designed antenna structure is given in Fig. 1 with its top and bottom views. Grey parts represent the copper that has a thickness of 0.035 mm and the white sections are the substrate. Table 1 presents the dimensional values of the designed antenna structure. Rogers RT5880 material is used as a substrate with a thickness of 1.575 mm. It has a dielectric constant of 2.2 and a loss tangent of 0.0009. The presented antenna element has a total size of 50 mm×43 mm×1.645 mm.

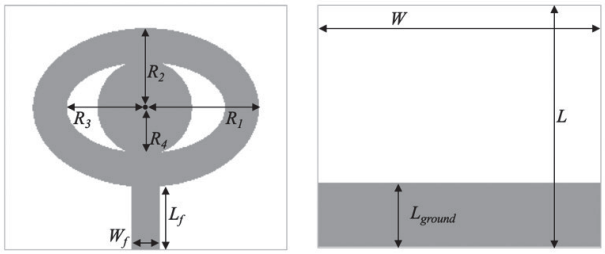


Figure 1: Top and bottom view of the designed antenna structure.

Table 1: Antenna dimension parameters.

Parameter	*W*	*L*	L_{ground}	L_f	W_f	R_1	R_2	R_3	R_4
Value (mm)	50	43	11.5	11	5	20	14	14	8.2

A parametric study is performed to justify the optimum antenna dimensions in Fig. 2, where the return loss (S_{11}) versus frequency graphs are plotted for analyzes. The width of the feeding line is varied from 4 to 6 in Fig. 2(a). As seen from the graph, increasing the width shifts the graph to the left and the optimum result is obtained with 5 mm value. In Fig. 2(b) the ground plane length is analyzed. It is observed that increasing the ground plane length shifts the curve to the right and more importantly, the widest bandwidth is achieved with the value of 11.5 mm. Another parametric study is simulated for substrate thickness and its effect on return loss are shared in Fig. 2(c). RT5880 material is available with 1.575 mm and 3.175 mm values. When both are compared, it is observed that the sub-6 GHz bands are covered with 1.575 value. On the other hand, a dual narrow band characteristic is appeared with 3.175. The return loss graph for different substrate materials are plotted in Fig. 2(d). The default RT5880 material is compared with FR-4, glass and RO4003 materials with their available thicknesses in market. Again, the best and desired result is obtained with the default RT5880.

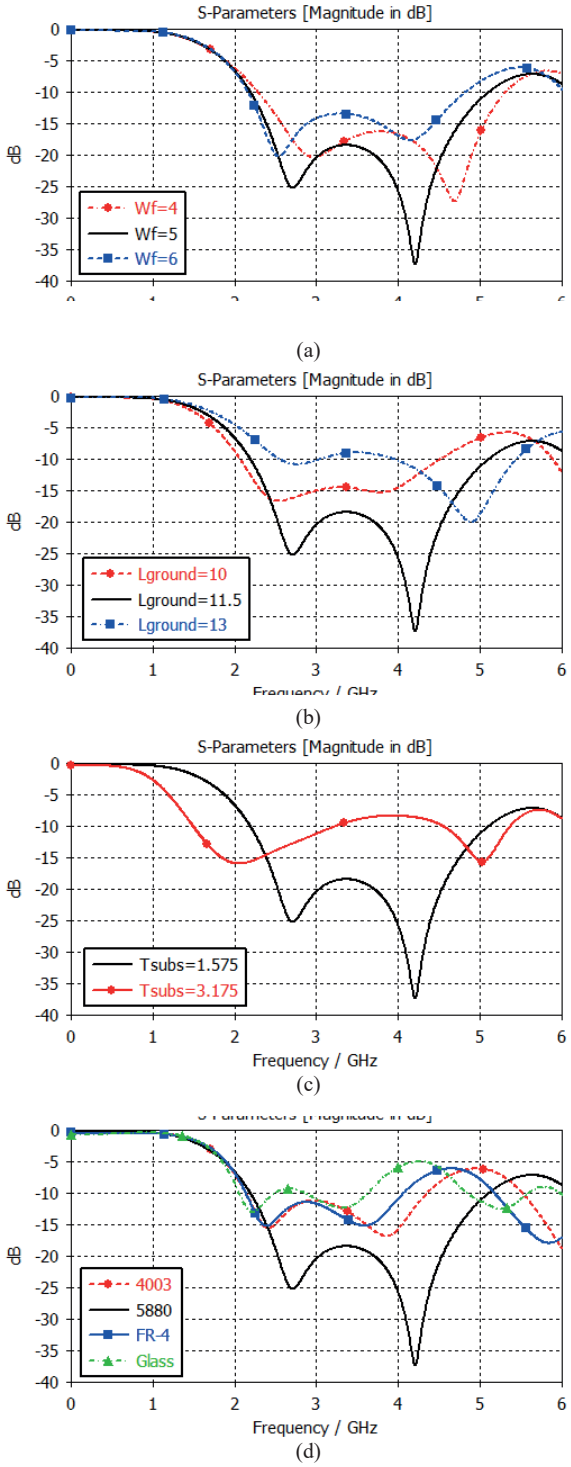


Figure 2: Return loss versus frequency curves for different (a) feed line widths, (b) ground lengths, (c) substrate thicknesses and (d) substrate materials.

III. Simulation Results

The simulation results of the antenna are performed with the help of CST Microwave Studio 2018 [13]. The first analysis is performed for return loss. Fig. 3 depicts the frequency versus return loss graph. It is observed from the graph that the antenna has an impedance bandwidth ($S_{11} \leq -10\,dB$) of 2.9 GHz between 2.195 GHz to 5.095 GHz. Such an operating frequency band covers all three previously mentioned sub-6 GHz bands. Furthermore, the proposed antenna has a fractional bandwidth percentage of 80%, which is an excellent result once compared with the values in [1-3] and [5-11]. All these outputs confirm the single wideband characteristic of the antenna. In addition, the antenna has a resonance at 4.2 GHz frequency, where S_{11} value decreases down to -37 dB.

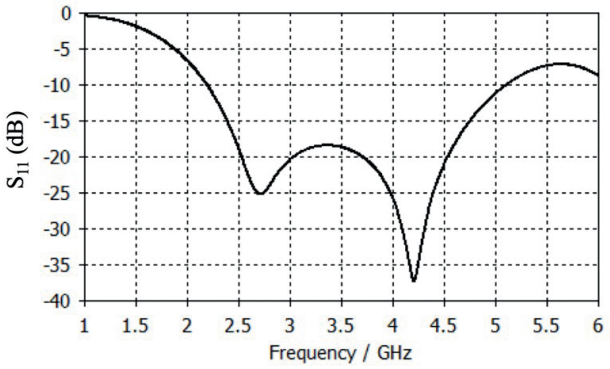


Figure 3: Return loss versus frequency graph.

The radiation efficiency and realized gain graphs are plotted in Fig 4 and Fig. 5, respectively. As seen from Fig. 4, the efficiency of the single antenna element is well above 87% in the operating frequency range. It has its maximum value of 96% at 4.6 GHz. This is a high efficiency result in comparison to other single element antenna designs. Moderate gain results are obtained from the simulation with the maximum value of 3.2 dBi at 5.05 GHz. Furthermore, one can say that the average gain value throughout the 2.195 to 5.095 GHz range is around 2.25 dBi, which is acceptable for an omni-directional characteristic antenna.

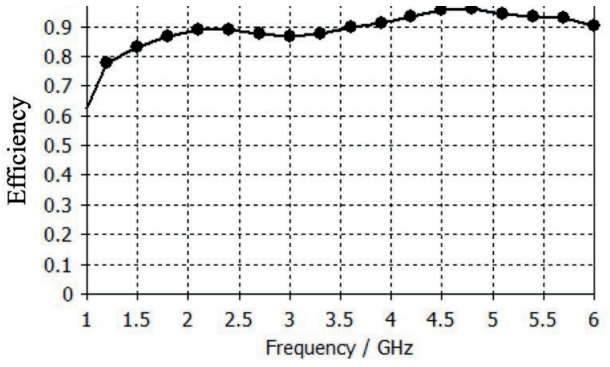


Figure 4: Radiation efficiency of the antenna.

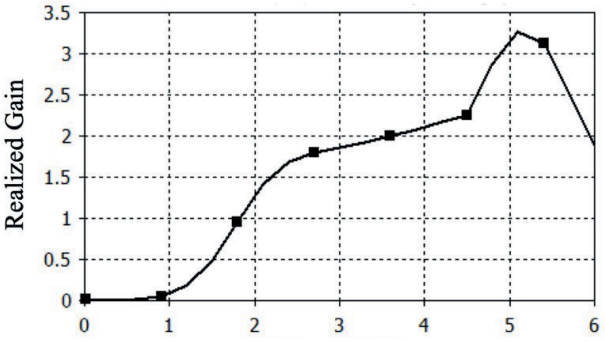


Figure 5: Realized gain of the antenna.

Radiation pattern of the antenna also needs to be investigated for a proper analysis. Fig. 6 represent the radiation pattern in polar form for *E* and *H*-planes. As seen from the graphs the antenna has an omni-directional radiation characteristic at the resonance frequency of 4.2 GHz, especially for its *E*-plane. This indicates that it nearly transmits/receives equally in all directions.

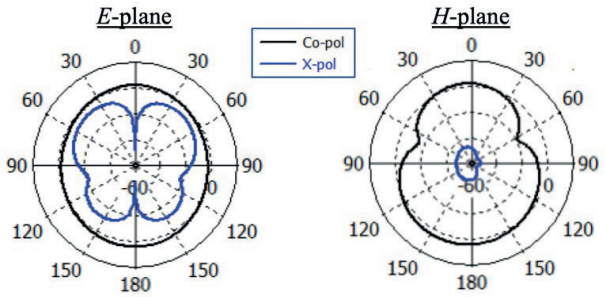


Figure 6: Radiation pattern at the resonance frequency.

IV. MIMO Antenna

MIMO antennas are frequently implemented in wireless communication systems thanks to their ability in increasing data capacity. The 4×4 MIMO antenna configuration of the proposed design is given in Fig. 7. As seen from the figure, the four antenna elements are positioned perpendicular to each other to prevent mutual coupling between them. Moreover, the MIMO antenna has a size of 73 mm×73 mm and it is designed by using the same 1.575 mm thick RT5880 substrate.

Figure 7: 4×4 MIMO antenna structure.

S-parameters are plotted in Fig. 8 for 4×4 MIMO structure. The return loss (S_{11}, S_{22}, S_{33}, S_{44}) and insertion loss (S_{12}, S_{13}, S_{14}, S_{21}, S_{23}, S_{24}, S_{31}, S_{32}, S_{34}, S_{41}, S_{42}, S_{43}) curves are generally used to determine the operating bandwidth and isolation between the antennas, respectively. As seen from the figure, the isolation between the antennas are below -18 dB within the operating frequency range. It indicates that the isolation between antennas are maintained successfully.

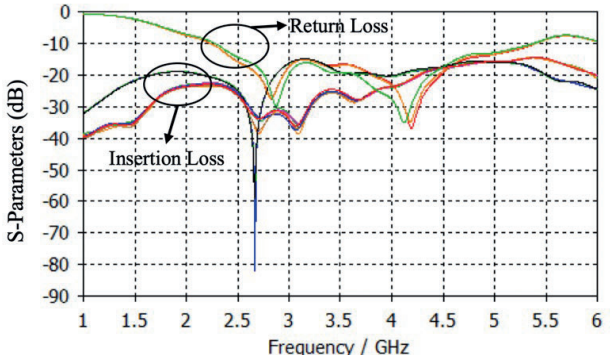


Figure 8: S-parameter analysis of MIMO antenna.

V. Conclusion

An eye shaped microstrip fed antenna is proposed in this study that has an operating bandwidth from 2.195 GHz to 5.095 GHz (80%). Therefore, it covers the all three sub-6 GHz 5G bands. Beside its wide bandwidth, it also has a high maximum radiation efficiency of 96%. As a future work, it is planned to fabricate the antenna and perform its experimental measurements with a vector network analyzer to properly validate its compatibility.

References

[1] M I Khattak, A Sohail, U Khan, Z Barki, and G Witjaksono, "Elliptical Slot Circular Patch Antenna Array with Dual Band Behaviour for Future 5G Mobile Communication Networks," *Progress In Electromagnetics Research C*, vol 89, pp 133–147, 2019

[2] Y Li, Z Zhao, Z Tang, and Y Yin, "A low-profile, dual-band filtering Antenna with high selectivity for 5G sub-6 GHz applications," *Microwave and Optical Technology Letters*, vol 61, no 10, pp 2282-2287, 2019

[3] A K Saurabh, and M K Meshram, "Compact sub-6 GHz 5G multiple-input-multiple-output antenna system with enhanced isolation," *International Journal of RF and Microwave Computer-Aided Engineering*, vol 30, no 8, e22246, 2020

[4] T Okan, "A compact octagonal-ring monopole antenna for super wideband applications," *Microwave and Optical Technology Letters*, vol 62, no 3, pp 1237-1244, 2020

[5] A Desai, T Upadhyaya, J Patel, R Patel, and M Palandoken, "Flexible CPW fed transparent antenna for WLAN and sub-6 GHz 5G applications," *Microwave and Optical Technology Letters*, vol 62, no 5, pp 2090-2103, 2020

[6] S S Singhwal, B K Kanaujia, A Singh, J Kishor, and L Matekovits, "Dual-band circularly polarized MIMO DRA for sub-6 GHz applications," *International Journal of RF and Microwave Computer-Aided Engineering*, vol 30, no 10, e22350, 2020

[7] A Zaidi, W A Awan, N Hussain, et al , "A Wide and Tri-band Flexible Antennas with Independently Controllable Notch Bands for Sub-6-GHz Communication System," *Radioengineering*, vol 29, no 1, pp 44-51, 2020

[8] M Mantash, A Kesavan and T A Denidni, "Beam-Tilting Endfire Antenna Using a Single-Layer FSS for 5G Communication Networks," *IEEE Antennas and Wireless Propagation Letters*, vol 17, no 1, pp 29-33, 2018

[9] T Okan, "Design and analysis of a quad-band substrate-integrated-waveguide cavity backed slot antenna for 5G applications," *International Journal of RF and Microwave Computer-Aided Engineering*, vol 30, no 7, e22236, 2020

[10] S Li, T Chi, Y Wang and H Wang, "A Millimeter-Wave Dual-Feed Square Loop Antenna for 5G Communications," *IEEE Transactions on Antennas and Propagation*, vol 65, no 12, pp 6317-6328, 2017

[11] J Park, J Ko, H Kwon, B Kang, B Park and D Kim, "A Tilted Combined Beam Antenna for 5G Communications Using a 28-GHz Band," *IEEE Antennas and Wireless Propagation Letters*, vol 15, pp 1685-1688, 2016

[12] C A Balanis, Antenna theory: analysis and design New York: J Wiley; 2005

[13] CST Microwave Studio Suite 2018, CST Inc

A Passive Sensor Design for Structural Health Monitoring by using Frequency Selective Surfaces

B. DÖKEN[1]

[1] Istanbul Technical University, Istanbul/Turkey, dokenb@itu.edu.tr

Abstract **- Structural health monitoring (SHM) aims to ensure detection and prevention of damage such as tilting, crack, movement, etc. in civil structures. Bridges, buildings, highways, tunnels, power plants, and dams are among the application areas of SHM. In this work, a novel frequency selective surface (FSS) geometry which acts as a passive sensor for tilt angle detection at civil structures is presented. FSSs are periodic conductive geometries which act as spatial filters. The proposed FSS sensor can be mounted on the exterior and interior surfaces of the buildings. The tilt angle of civil buildings can be detected from 10 to 60 degrees by using proposed FSS sensor.**

Keywords **- Absorber, Frequency Selective Surface, FSS, Periodic Structures, Structural Health Monitoring, SHM, Tilting**

I. INTRODUCTION

The structural health monitoring (SHM) systems are installed on bridges, building, highways, tunnels, power plants, and dams to improve the capacity of damage detection [1]–[3] . Many types of sensors have been developed to monitor structural health conditions. These sensors are divided into two types: active and passive sensors. Active sensors are commonly used for monitoring system in SHM to detect any movement or structural changes such as stress and tension. By using active sensors, power consumption and high cost for maintenance is required to perpetuate the performance. Passive sensors have the advantage of easier installation and maintenance compared to active sensors, and no power consumption.

FSSs are periodic conductive geometries which behave as a frequency selective filter [4]–[7]. Their frequency behavior depends on their geometries, the parameters of these geometries, the properties of the dielectric layers on which they are mounted, and the frequency, polarization and incidence angle of the incoming electromagnetic waves. Antennas, reduction of electromagnetic interference, absorber surfaces are some of the application areas of FSSs. The use of FSS as a sensor in SHM is another important research topic [8]–[10]. FSS sensors are mounted to the surface of existing structures in order to monitor normal and shear strain damages, or in case of layered structures to detect delamination. Since FSS response depends on its geometry and to the incidence angle of the incoming electromagnetic wave, any deformation at civil structure leads to a change on FSS's frequency response. By measuring the deviations in frequency response of FSS, the damage or deformation of civil structures can be calculated.

This work proposes an FSS-based passive sensor for SHM, which is mounted on the surfaces of the civil structures. The ground of the FSS sensor is covered with a conductive layer to isolate sensor from the civil surface. Therefore, the frequency response of the FSS sensor is stable against to the changes in the electrical properties of the surfaces of civil structures. Proposed FSS sensor geometry has highly sensitive to incidence angle response and tilting angle of civil buildings can be detected from 10 to 60 degrees.

The use of FSSs in SHM is a new research topic. 3D FSS geometries are proposed as a passive sensor with high sensitivity to incident angles to be apply in SHM system [2] and [8]. In [11] a new application of FSSs is discussed as a potential inspection method for detection of delamination in layered structures. A strain sensor which is isolated from the surfaces of civil structures is presented in [9] by using FSS geometries. To the best of the authors knowledge, except this work, there is no tilting FSS sensor in the literature which isolates FSS structure from the surfaces of civil structures.

This paper is organized as follows. Section II describes the design steps and presents the simulation results. Section III discusses the results.

II. DESIGN & SIMULATION

The aim of this work is to design an FSS sensor for detection the tilting angle at buildings, bridges, dams, remotely. The resonance frequency of FSS depends on the incident angle of the incoming electromagnetic wave [4]. As shown in Figure 1, the angle of incidence angle of FSS is same to the angle of tilting of building. The FSS sensor to be designed is on the exterior or interior surfaces of the buildings. Therefore, the occurred tilt angle at civil structures leads to change on frequency response of FSS. In this work, a passive SHM sensor design is utilized by using this feature of FSS.

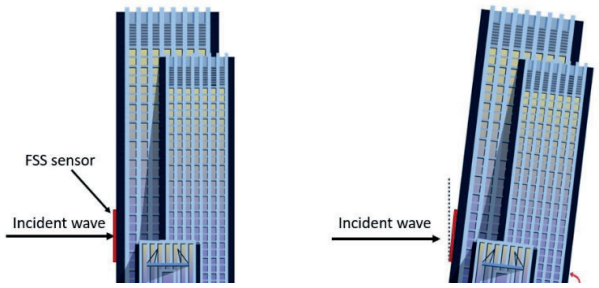


Figure 1: Tilted building

At the first design stage, “Square Loop” element geometry (Figure.2a) is selected as the initial FSS geometry due to having polarization insensitivity and simple in structure. FSS sensor (Figure.2b) is then formed by printing “Square Loop” FSS geometry on the other side of the 1.6mm FR4 (ϵ_r=4.4) substrate coated with conductor. 1.6 mm FR4 substrate is selected to easily perform prototype fabrication at the design.

The electrical properties (dielectric permittivity, loss tangent, etc.) of the civil surfaces depend on parameters such as cement density, moisture content etc. The deviations on these parameters lead to changes on the frequency response of FSS sensor. The frequency response of the FSS sensor must be stable to the changes in the electrical properties of the surfaces of civil structures. Therefore, in this work, the ground of the FSS sensor is covered with a conductive layer to isolate sensor from the surface.

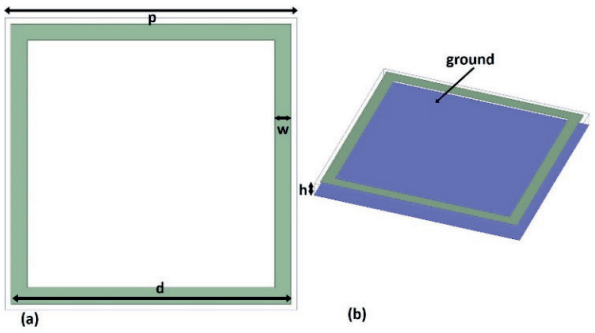


Figure 2: Unit cell and parameters of the FSS: (mm) p=25, d=24, w=1.4, and h=2 (a) Top view (b) Perspective view.

Transmission parameter (S_{21}) of the FSS sensor is equal zero due to the existing of conductive ground layer. The surface impedance of FSS sensor (Z_{in}) depends on its geometry and its parameters and to the FR4 substrate itself [4], [12]. According to Eq.1 [13], reflection coefficient is controlled by the FSS sensor geometry and by the FR4 substrate itself. Z_0 denotes the free space impedance of the vacuum in Eq.1. Therefore, by a proper optimization stage, surface impedance of the proposed FSS is optimized to have a resonance frequency between 1 GHz to 2 GHz frequencies.

$$S_{11} = (Z_{in} - Z_0)/(Z_{in} + Z_0) \quad (1)$$

The tilting angle of civil structure leads to a proportional (linear) change at resonance frequency of FSS sensor. The "Square Loop" geometry has been chosen for its suitability for this purpose.

Analyses and optimization of the FSS sensor is performed by Ansoft HFSS 19.1 software. Parametric analysis feature of Ansoft HFSS software is used in the optimization stage. Figure 3 and Figure 4 shows the obtained reflection coefficient (S_{11}) results ranging from 0^0 to 50^0 of incidence angles for TE (Transverse Electric) and TM (Transverse Magnetic) polarizations, respectively.

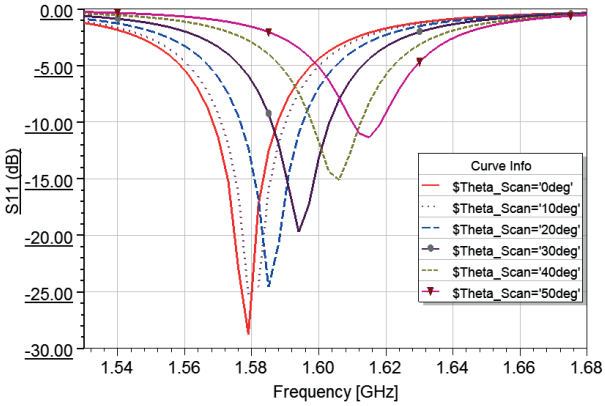


Figure 3: S_{11} simulation results (TE polarization)

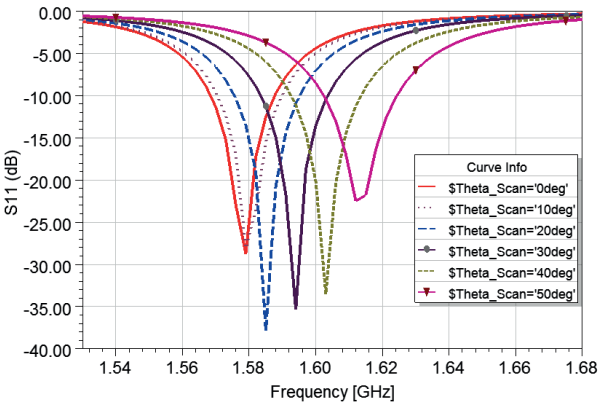


Figure 4: S_{11} simulation results (TM polarization)

According to the simulation results, minimum 20 dB reflection coefficient attenuation is achieved which is enough for measurement purposes. The resonance frequency deviation of the reflection coefficient with respect to the angle of incidence is seen in Figure 5 and Figure 6 for TE and TM polarizations, respectively.

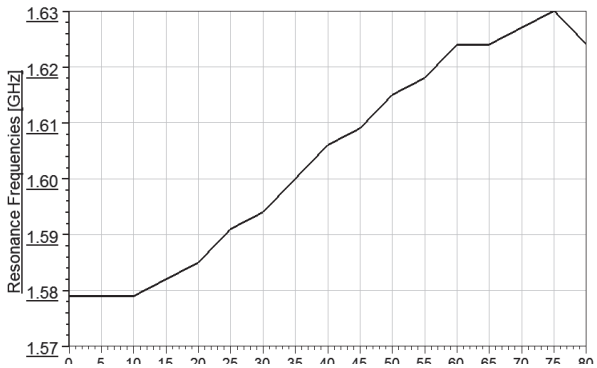


Figure 5: Resonance frequencies of S_{11} for different incidence angles at TE polarization

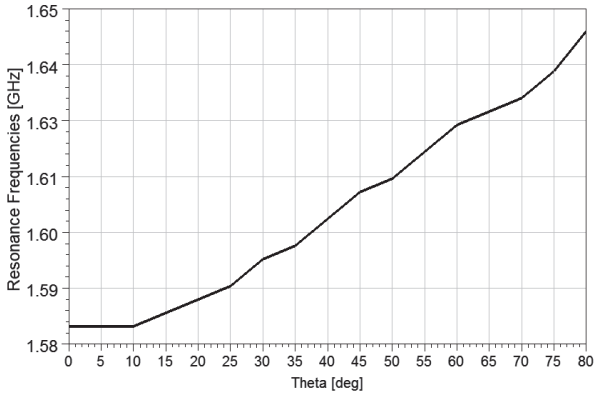


Figure 6: Resonance frequencies of S_{11} for different incidence angles at TM polarization

According to the simulation results, it is seen that the resonance frequency change of FSS is linear with the change of incidence angle. Therefore, the proposed FSS sensor can measure tilting angles between 10 degrees to 60 degrees.

III. Conclusion

The deviation of resonance frequency of the proposed FSS with respect to incidence angle is almost linear between 10^0 degrees to 60^0. The proposed design is also very thin (2mm) and can be easily mounted on structural surfaces. Tilting angle in structures can be easily detected remotely with the proposed design. However, as an important deficiency, tilting angles between 0 and 10 degrees cannot be determined.

References

[1] J M López-Higuera, L R Cobo, A Q Incera, and A Cobo, 'Fiber optic sensors in structural health monitoring', *J. Light. Technol.*, vol 29, no 4, pp 587–608, 2011

[2] S A Suhaimi, S N Azemi, P J Soh, C B M Rashidi, and A A Al-Hadi, 'Highly sensitive frequency selective surface for structural health monitoring system', *Indones. J. Electr. Eng. Comput. Sci.*, vol 14, no 2, pp 523–528, 2019

[3] Y Tan, J Zhu, and L Ren, 'A two-dimensional wireless and passive sensor for stress monitoring', *Sensors*, vol 19, no 1, p 135, 2019

[4] B A Munk, *Frequency Selective Surfaces - Theory and Design* New York: John Wiley and Sons Inc , 2000

[5] P C Zhao, Z Y Zong, W Wu, and D G Fang, 'A Convoluted Structure for Miniaturized Frequency Selective Surface and Its Equivalent Circuit for Optimization Design', *IEEE Trans. Antennas Propag.*, vol 64, no 7, pp 2963–2970, 2016

[6] A Using, L C Resonators, K Zhang, W Jiang, and S Gong, 'Design Bandpass Frequency Selective Surface', *IEEE Antennas Wirel. Propag. Lett.*, vol 16, pp 2586–2589, 2017

[7] J C Vardaxoglou, *Frequency selective surfaces: analysis and design* Research Studies Press, 1997

[8] S A Suhaimi, S N Azemi, and P J Soh, 'Feasibility study of frequency selective surfaces for structural health monitoring system', *Prog. Electromagn. Res. C*, vol 80, pp 199–209, 2018

[9] E Kinzel, 'Design of a Frequency-Selective Surface strain sensor', in *2014 IEEE Antennas and Propagation Society International Symposium (APSURSI)*, 2014, pp 2074–2075

[10] S A Suhaimi, S N Azemi, and S P Jack, 'Structural health monitoring system using 3d frequency selective surface', in *2016 IEEE Asia-Pacific Conference on Applied Electromagnetics (APACE)*, 2016, pp 145–149

[11] D F Pieper and K M Donnell, 'Application of frequency selective surfaces for inspection of layered structures', in *2015 IEEE International Instrumentation and Measurement Technology Conference (I2MTC) Proceedings*, 2015, pp 1204–1209

[12] T Liu and S S Kim, 'Design of wide-bandwidth electromagnetic wave absorbers using the inductance and capacitance of a square loop-frequency selective surface calculated from an equivalent circuit model', *Opt. Commun.*, 2016

[13] D M Pozar, *Microwave engineering* John Wiley & Sons, 2009

Memristor Based PD Controller Design and Application on a Ball and Beam Control System

K. ORMAN

Erzincan Binali Yıldırım University, Erzincan/Turkey, korman@erzincan.edu.tr

Abstract **- In this study, a memristor based PD (Mem-PD) controller is designed and the performance of the controller is tested in the simulation model of the beam and ball (Ball and Beam) control system. In the proposed control structure, the position of the ball in the primary part called the outer loop is controlled by changing the angle of the beam. In the secondary part, called the inner loop, voltage is generated that will determine the proper position angle of the servomotor to adjust the position of the beam. In addition, the use of Memristor, which has attracted increasing attention in recent years, in the field of control systems has been tested. The simulation results show that the designed controller successfully performed reference tracking.**

Keywords **– PD Control, Memristor, Ball and Beam, MemPD.**

I. INTRODUCTION

MEMRISTOR, first proposed by Leon Chua in 1971 as a lost circuit element, was defined as a memory resistor [1]. After the memristor was produced in HP laboratories in 2008, it became more popular and its popularity increased [2]. In order to create this circuit element, first, silicon dioxide or titanium dioxide is obtained by using titanium and silicon semiconducts and oxygen atoms, and by using any of these compounds, the internal structure of the memristor is formed. As the current flows through the memristor, the oxygen spaces in the doped titanium dioxide layer slide towards the direction of the current, that is, pure titanium dioxide, as a result, the pure titanium dioxide range narrows. The spacing length of doped titanium dioxides increases and the boundary between TiO_2 film layers changes, causing the resistance to change. In the absence of current, the doped and pure TiO_2 layers maintain their gap levels. In the opposite case, that is, with the negative flow of current, the oxygen spaces move towards the doped layer. Then the pure TiO_2 layer increases its width, causing an increase in resistance. Until today, many studies have been carried out in different fields by using memristor models and emulator circuits [3-9]. In addition, studies were using memristor in the field of control systems have become widespread in recent years. Delgado stated in his study that the Memristor can be used as an adjustable gain for closed loop systems. Also it has been stated that the memristor acts as a linear resistor at high frequencies [10]. Wnag et al. [11] presented a mathematical model of an M-PID controller by replacing a resistance of a conventional PID control circuit with a memistor. They stated that the M-PID controller behaves faster and stronger in terms of dynamic performance compared to the traditional PID device. Wang et al. [12] addressed the problem of finite-time stabilization of memistor-based delayed neural networks (MDNNs) through two control approaches in their study. Lu et al. [13] presented self-adjusting PID circuits with voltage and current controlled memristor in their work. In addition, simulation and experimental results are presented to show that the structure with memristor has a more suitable robustness than traditional PID controller. In a study in which parameter update rules of voltage and current controlled memristors were discussed [14]; The performance of the closed loop system (DC / DC power convertor) was tested under the control of the memristor based adaptive PI controller. With simulations and experimental results, it is stated that the controller parameters can vary with the operating point of the controlled system compared to the traditional PI controller, so the proposed controller is more robust and is suitable for scenarios with large power supplies and load disturbances. In another controller design application made with memristor [15]; Simulation study of temperature profile tracking control in heat flow system using memristor based PI controller is presented. The success of the Mem-PI controller, which is obtained by using memristor instead of the gain resistor in traditional PI controller, is presented with simulation results.

Ball and Beam experimental setup is a system widely used in many university laboratories to show students the applications of control algorithms [16-19]. The purpose of this system is to control the position of the ball on the beam.

In this study, in order to test the usability of the Memristor in the field of control systems, the traditional PD controller was redesigned using a memristor and the Mem-PD controller obtained was applied to the Ball and Beam system. Simulation results are presented to demonstrate the recommended controller success.

II. MODELLING OF BALL AND BEAM SYSTEM

The ball and beam experimental setup system was designed by Quanser, and information about the system is clearly described in their documentation [20]. The schematic diagram of the ball and beam system is shown in Fig. 1. Here m_b is the mass of the ball, r_b is the radius of the ball, γ_b is the angle of the ball, x is the position of the ball, α is the angle with the beam axis, g_b is gravitational acceleration and J_b is momentum.

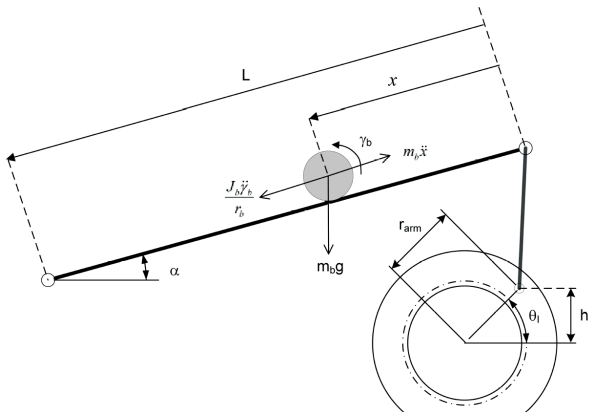


Figure 1: Free body diagram of the Ball and Beam system.

In the light of Ref. [20] the linearized acceleration equation of the ball can be obtained as follows.

$$\ddot{x}(t) = \frac{m_b g \sin\alpha(\mathrm{t}) r_b^2}{m_b r_b^2 + J_b}. \tag{1}$$

The relationship between the angle of the servo motor $\theta_l(t)$ and the length of the beam L_{beam} can be expressed as follows, considering the servo dynamics.

$$\sin\alpha(t) = \frac{\sin\theta_l(t) r_{arm}}{L_{beam}}. \tag{2}$$

Here, the r_{arm} is the distance between the plated joint and the output gear shaft. The relation mentioned in (2) can be added to the nonlinear motion equation found in [20]. The final form of the dynamic equation of the ball beam system after mathematical manipulation is as follows.

$$\ddot{x}(t) = \frac{m_b g \theta_l(t) r_{arm} r_b^2}{L_{beam}(m_b r_b^2 + J_b)}. \tag{3}$$

III. Controller Design

If x_d is the desired position and x is the current position, position tracking error of the Ball and Beam system for the controller design can be defined as follows.

$$e(t) = x(t) - x_d(t). \tag{4}$$

The PD controller designed using op-amp is shown in Fig. 3. The capacitor in the circuit is used to store charge and represents the differential equation of the input. In the PD controller circuit, the output $u(t)$ can be written as below;

$$u(t) = -\frac{R_2}{R_1} e(t) - R_2 C \frac{d}{dt} e(t). \tag{5}$$

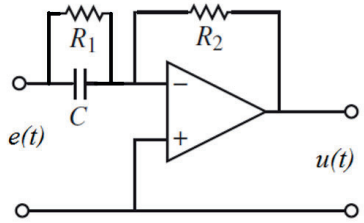


Figure 2: Schematic diagram for PD controller using operational amplifier.

The following statement can express the relationship between the current and voltage of the Memistor.

$$V(t) = R(t) i(t) = \frac{d\varphi}{dQ} i(t). \tag{6}$$

Here, Q and φ denote charge and flux, respectively. The ratio of flux to charge is called memristance and refers to the resistance of the memristor. The relationship between current and voltage in TiO_2 memistor has been proposed by the HP research group [7] as follows.

$$V(t) = \left[R_{ON} \frac{X(t)}{D} + R_{OFF}\left(1 - \frac{X(t)}{D}\right) \right] i(t). \tag{7}$$

In the above equation, R_{ON} and R_{OFF} refer to high and low resistance values, respectively. X and D indicate and doped area and the thickness of the TiO_2 memristor, respectively. Resistance value, which depends on the direct change of the X value, is defined as:

$$\frac{dX(t)}{dt} = \mu_v \frac{R_{ON}}{D} i(t). \tag{8}$$

Here μ_v refers to the mobility of the memistor. Memistors show a stuck hysteresis voltage-current relationship when a sinusoidal signal is applied, and there is an inverse ratio between frequency and nonlinearity as shown in Fig. 3.

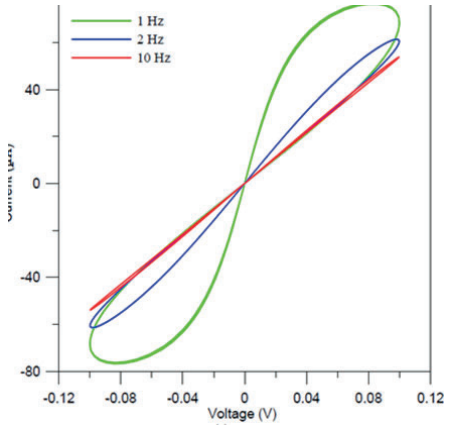


Figure 3: Memistor current-voltage relationship for various frequencies [7].

To achieve the Mem-PD controller, we can replace R_2 with

a memistor in the circuit given in Fig. 2. In this way, the circuit belonging to the Mem-PD controller will be as in the Fig. 4 below.

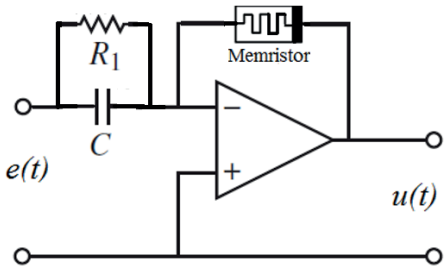


Figure 4: The Mem-PD controller circuit

Considering (5), (6) and (7), if the control signal *u(t)* in the circuit given in Fig. 4 is rewritten;

$$u(t) = -\frac{\left[R_{ON}\frac{X(t)}{D} + R_{OFF}(1-\frac{X(t)}{D})\right]}{R_1}e(t) - \left[R_{ON}\frac{X(t)}{D} + R_{OFF}(1-\frac{X(t)}{D})\right]C\frac{d}{dt}e(t). \quad (9)$$

Considering Ref. [20], the block diagram of the position control of the Ball and Beam system with Mem-PD control approach is given in Figure 5.

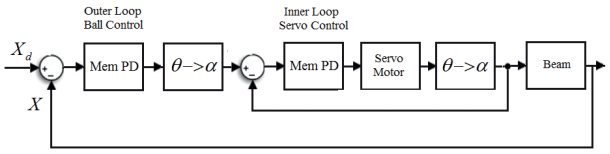


Figure 5: Position control of the Ball and Beam system using Mem-PD controller

IV. Simulation

In this section, to test the performance of both the conventional PD controller and the designed Mem-PD controller, a square wave position reference signal with sudden changes was applied to the ball and beam system and the simulation results were obtained and given in Fig. 6-8.

When the Fig. 6 showing the position of the ball is examined, it is observed that both controllers show approximately the same performance. However, according to the zoomed image, it is seen that the Mem-PD controller reached the reference 0.4 seconds earlier than the PD controller.

In the graph showing the servo angles and given in Fig. 7, especially during the sudden changes in the 5th and 10th seconds, the PD controller has 3% more overshot than the Mem-PD controller.

Fig. 8 shows the control signals of both control devices. Although similar signals are generated, the amplitude of the control signal of the PD controller is 2% more than the control signal amplitude produced by the Mem-PD at the sudden change moments.

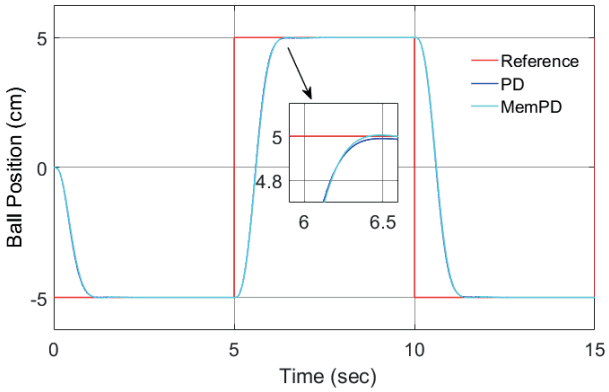


Figure 6: Simulation result of Ball positon of the Ball and Beam system using PD and Mem-PD controllers

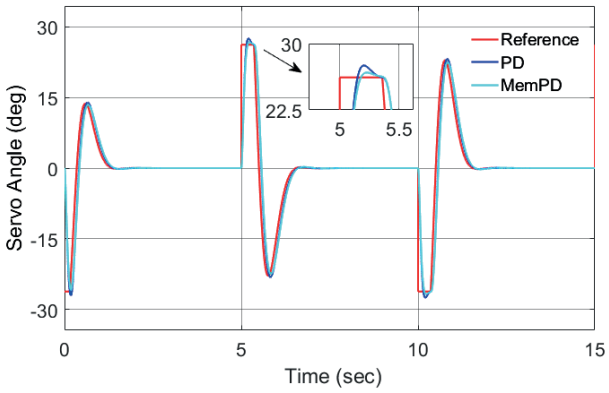


Figure 7: Servo angles for PD and Mem-PD controller

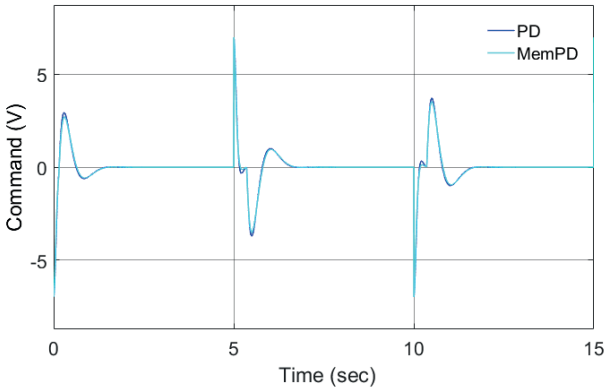


Figure 8: Command signals for PD and Mem-PD controllers

V. Conclusion

In this study, a memristor based PD controller is proposed. The control structure was achieved by using a memistor instead of the gain resistor in the traditional PD controller circuit. The performance of the designed Mem-PD controller was compared with the traditional PD controller, and the simulation results showed that the proposed controller was successful in both reference tracking and error recovery. In

future work, it is planned to test this controller in a real time experimental setup.

REFERENCES

[1] Chua, L "Memristor-the missing circuit element," *IEEE Transactions on circuit theory*, 18(5), 507-519 1971

[2] Strukov, D B , Snider, G S , Stewart, D R , & Williams, R S "The missing memristor found," *nature*, 453(7191), 80-83 2008

[3] Muthuswamy, B , & Kokate, P P "Memristor-based chaotic circuits," *IETE Technical Review*, 26(6), 417-429 2009

[4] Jo, S H , Chang, T , Ebong, I , Bhadviya, B B , Mazumder, P , & Lu, W "Nanoscale memristor device as synapse in neuromorphic systems," *Nano letters*, 10(4), 1297-1301 2010

[5] Thomas, A "Memristor-based neural networks," Journal of Physics D: *Applied Physics*, 46(9), 093001 2013

[6] Yeşil, A , Babacan, Y , & Kaçar, F "A new DDCC based memristor emulator circuit and its applications," *Microelectronics Journal*, 45(3), 282-287 2014

[7] Babacan, Y , & Kaçar, F "Memristor emulator with spike-timing-dependent-plasticity," *AEU-International Journal of Electronics and Communications*, 73, 16-22 2017

[8] Zidan, M A , Jeong, Y , Lee, J , Chen, B , Huang, S , Kushner, M J , & Lu, W D "A general memristor-based partial differential equation solver," *Nature Electronics*, 1(7), 411-420 2018

[9] Yeşil, A , Babacan, Y , & Kaçar, F "Design and experimental evolution of memristor with only one VDTA and one capacitor," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 38(6), 1123-1132 2018

[10] Delgado, A "The memristor as controller," *IEEE Nanotechnology Materials and Devices Conference* (pp 376-379) IEEE October 2010

[11] Wang, X , Zhao, Y , & Liao, Y "Dynamic performance analysis of PID controller with one memristor," *International Conference on Information Science and Technology* (pp 1234-1237) IEEE March 2011

[12] Wang, L , Shen, Y , & Zhang, G "Finite-time stabilization and adaptive control of memristor-based delayed neural networks," *IEEE transactions on neural networks and learning systems*, 28(11), 2648-2659 2016

[13] Lu, Y , Liang, Q , & Huang, X "Parameters self-tuning PID controller circuit with memristors," *International Journal of Circuit Theory and Applications*, 46(1), 138-154 2018

[14] Lu, Y , Wang, Y , & Huang, X "Adaptive memristor-based PI control of a DC/DC converter non-minimum phase system," *IEEE International Power Electronics and Application Conference and Exposition* (PEAC) (pp 1-6) IEEE November 2018

[15] Orman, K "Temperature Profile Tracking Control with Memristor Based PI Controller in the Heat Flow System," *International Journal of Modern Research in Engineering and Technology* 4(12), pp 12-15 2019

[16] Hauser, J , Sastry, S , & Kokotovic, P "Nonlinear control via approximate input-output linearization: The ball and beam example," *IEEE transactions on automatic control*, 37(3), 392-398 1992

[17] Yu, W "Nonlinear PD regulation for ball and beam system," *International Journal of Electrical Engineering Education*, 46(1), 59-73 2009

[18] Maalini, P M , Prabhakar, G , & Selvaperumal, S "Modelling and control of ball and beam system using PID controller," *International Conference on Advanced Communication Control and Computing Technologies* (ICACCCT) (pp 322-326) IEEE May 2016

[19] Mehedi, I M , Al-Saggaf, U M , Mansouri, R , & Bettayeb, M "Two degrees of freedom fractional controller design: Application to the ball and beam system," *Measurement*, 135, 13-22 2019

[20] Quanser Innovative Edutech, Ball and Beam Laboratory Manual, Quanser Inc , 2012

An Experimental Setup for Determining Optimum Tilt and Orientation in Photovoltaic Panels

M.ARSLAN[1] and M. ÇUNKAŞ[2]

[1] Konya Technical University, Vocational High School, Konya/Turkey, marslan@ktun.edu.tr
[2]Selcuk University, Faculty of Technology, Department of Electrical and Electronics Engineering, 42031, mcunkas@selcuk.edu.tr

***Abstract* - In Turkey, where annual solar energy potential is 380GWh, the amount of electrical energy that can be obtained by placing PV panels with optimum tilt angle is around 1.239kwh / 1kwp. One of the most effective methods of obtaining maximum level of electrical energy from PV panels is determining the optimum tilt angles and orientation angles of PV panels.**

In order to obtain the highest possible amount of electrical energy from PV panels for Konya province, two experimental setups, one of them is fixed and the other one is mobile, were established to determine the most appropriate Azimuth and Tilt angles. While the tilt of the panels can be changed from 0^0 to 60^0 with 15 degrees intervals, the azimuth angle can be changed by 10 degrees intervals. In order to calculate the electrical energy produced by the panels, current and voltage data is instantaneously collected from the experimental setup via invertors. Relationships between power outputs of PV panels according to the tilt and azimuth angle can be examined. In addition, the global radiation value is measured with the Pyranometer installed on the contraption. Moreover, ambient and panel temperatures are recorded simultaneously. The measurements are recorded with the help of a Data Acquisition card.

***Keywords* – PV Panel, Tilt Angle, Azimut Angle, Optimum Position**

I. INTRODUCTION

Energy consumption is increasing day by day due to technological developments in the world and Renewable Energy Resources are used to meet the demand. The common feature of all systems established to benefit from solar energy is based on making maximum use of the radiant energy of the sun. In photovoltaic applications all over the world, fixed mounted PV systems are installed in accordance with the latitude angle of the current location. In such applications, the total radiation values on the horizontal surface are taken into consideration.

When the potential solar energy atlas published by the General Directorate of Renewable Energy is examined, it is seen that Konya is the leader of the places with the highest solar radiation value. Konya is geographically located between 36.41' and 39.16' north latitudes and 31.14' and 34.26' east longitudes. Konya's solar energy potential reaches to the values above the averages of the year, especially in the spring and summer months. One of the important parameters affecting the performance of a solar panel is the tilt angle to the horizontal, this is because changing the angle of tilt changes the amount of sunlight reaching the panel surface.

There are many studies on solar systems in the literature. A mathematical model is used to determine the optimum tilt angle and direction (surface azimuth angle) of the solar collector in Brunei Darussalam. It reveals that rotating the tilt angle 12 times a year (by using the monthly average optimum tilt angle) maintains the total amount of solar radiation approximately same as the maximum value obtained by rotating the tilt angle daily to its optimum value [1]. For five cities in different geographical locations in the United States, the Perez radiation model of solar radiation coming to inclined surfaces is simulated by fixed, azimuth tracking and biaxial tracking of the PV panels. It has been observed that higher radiance values are obtained when single axis azimuth tracking is performed instead of a fixed panel facing south at optimum tilt angle. They found that adding a second tracking axis for tilt tracking in addition to the azimuth axis resulted in a 5% increase in irradiation compared to azimuth tracking [2]. It is aimed to determine the optimum tilt angle for PV panels in the Athens region, especially in summer. Experimental research was conducted for the entire summer period. In this context, whether the tilt angle of the panel differs from the expected optimum value, the study of its effect is emphasized, and accordingly, the results obtained are interpreted on the basis of the determined solar geometry equations [3].

A mathematical model was used to estimate the optimum tilt angle required for mounting the PV panel and to determine the total solar radiation on the tilted PV surface. Authors estimated the optimum tilt angles by studying the total radiation values on the PV surface. They also compared the energy gain of a fixed PV system and a two-axis mobile system and determined the optimum tilt angles for Şanlıurfa [4]. A study to determine the optimal tilt angle measurement data using solar radiation was carried out in major cities of Turkey. The optimum tilt angle was calculated from 0º to 90º using the periodic daily total solar radiation data on inclined surfaces in 10º intervals. It has been observed that the optimum panel tilt angle for these provinces throughout the year varies between 0º and 60º [5]. Using the diffused

radiation models given in the literature for Şanlıurfa, the diffused radiation values in the horizontal plane were calculated and the results obtained from the models were compared with the measurement data. Hourly horizontal total and diffused solar radiation data collected from 7-year measurement results obtained from solar tracking system and solar radiation measurement system in Şanlıurfa Harran University Mechanical Engineering Department were used. [6].

As can be seen when the studies on the calculations of the radiation on the tilt are examined, one of the most effective methods of obtaining the maximum annual energy from the PV plant is possible by determining the optimum tilt angles and orientation (azimuth) angles of the PV panels well. Moreover, it is necessary to know the solar radiation data of that region during the design and installation process of solar energy systems [7]. The tilt angle of a solar energy system is one of the most important parameters to capture the maximum solar radiation falling on solar panels [8]. Solar panel tilt angle is location-specific as it depends on the daily, monthly and yearly displacement of the sun. In calculating the tilt and orientation angles of PV panels, monthly average values of daily solar radiation on the horizontal plane are generally used and these data are available for many places around the world. However, radiation data, usually on inclined surfaces, are not available [9]. Therefore, mathematical models have been developed to predict irradiance data on inclined surfaces [10]. When installing solar PV systems, panels are positioned facing south and equator in the northern hemisphere, and facing north in the southern hemisphere [11]. In this way, the maximum level of solar radiation is collected on the panels [12]. In experimental studies, while determining the optimal tilt angle for PV modules, the panels were generally placed facing south direction for the northern hemisphere [13].

In this study, an experimental setup was established in order to form a solar radiation model specific to Konya by using both the radiation values coming to the horizontal plane and the radiation values coming to the inclined surface. Two experimental setups, one of which is fixed and the other one is mobile, were assembled. 310-Watt panels with 5 monocrystalline cell structures were used in the experimental setup. The generated electrical energy is transferred to the network by Micro inverters. Radiation pyranometers are mounted on the setup at the same tilt as the panels. Panel temperature sensors are mounted in the lower part of each panel. In addition, a temperature sensor is placed in the environment where the panels are located and the ambient temperature is recorded. Thanks to the experimental setup, not only the tilt angle of the PV module but also the its azimuth angle can be observed experimentally.

II. Material And Method

The tilt and orientation angles of PV panels used in the installation phase of solar power plants are important, because if the panels are not positioned at appropriate angles and directions, the energy production efficiency of the panel will be low since they will not be exposed to sufficient solar radiation. So, the amount of electrical energy obtained from the entire power plant will also be low. For this reason, in order to obtain the highest level of efficiency from PV panels, it is of great importance that they are placed at tilt and azimuth angles where they can receive higher radiation.

A. Experimental Setup

Two experimental setups, one of which is fixed and the other one is mobile, were assembled. As shown in Figure 1, the fixed panel was placed in the south direction at 0° azimuth angle and at a tilt angle between 36° and 39° (30°). These angles are for the latitude values where Konya is located.

Figure 1: Fixed PV panel

Figure 2: Movable PV panel to set tilt and azimuth angle

As can be seen in Figure 2, the other experimental setup was designed taking into account the compulsory distances between the panels and mechanical constraints to prevent cloudiness. It consists of 4 panels placed at different tilt angles on a structure where the azimuth angle can be adjusted in south-east and south-west directions. With this structure, it will be possible to obtain radiation values at 4 different tilts for the same azimuth angle. In the mobile system, the tilt angle of the PV panels is in range between 0° and 60° with a difference of 15° between each other. Azimuth angles can also be adjusted to be 0° to the south, 60° to the east and 60° degrees to the west. The azimuth angle increase rate is in the form of angle changes of 10°. As seen in Figure 3, the experimental test setup includes a radiation pyranometer with a module temperature sensor and an ambient temperature sensor to measure temperature values.

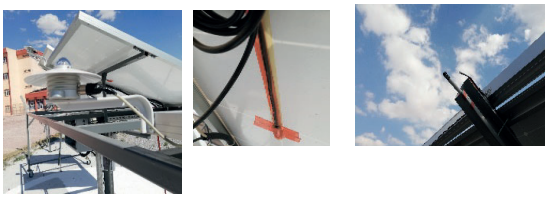


Figure 3: Placement of radiation pyranometers, panel temperature and ambient temperature sensors on the module

B. Data Acquisition and Analysis

The current, voltage and power values generated by each panel are recorded on the portal provided by the inverter company by means of micro inverters. In addition, a total of 5 radiation pyranometers were used, one on each inclined surface and one on the horizontal surface. Radiation and temperature values at constant tilt and 0º azimuth and power values under this irradiation are measured with the radiation pyranometer placed on the fixed mounting. The ambient temperature sensor placed in the horizontal plane measures the temperature of the environment. By placing the radiation pyranometer and module temperature sensor on all of the other four panels, it is possible to obtain values at different tilts with different azimuth angles. As seen in Figure 4, while micro-inverters transfer the collected data to computers, they also transmit the electrical energy generated by the PV panels to the grid. The system works as an on-grid.

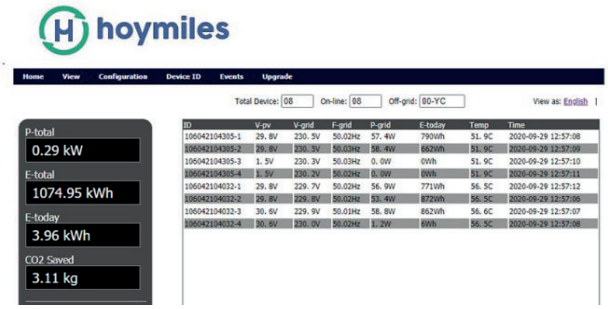
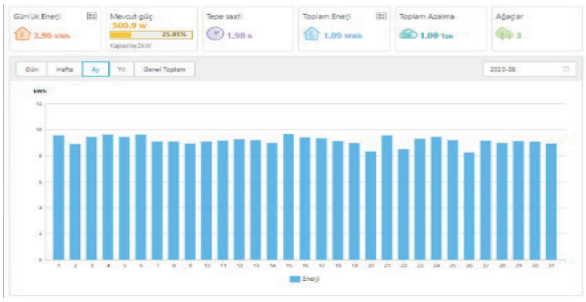


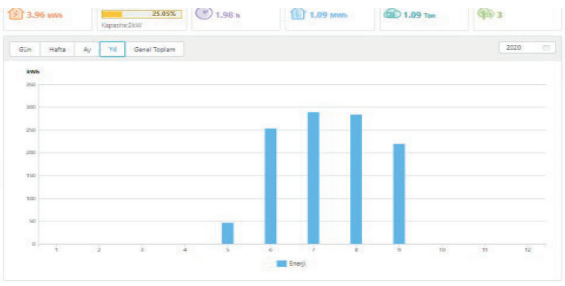
Figure 4: Micro-inverters and Data portal

The electrical energy generated by the PV panels is transferred to the city grid via micro-inverters. Inverters with dual MPPT monitoring, low voltage start and 16-60-volt operating range are used. Inverters also record the current, voltage and power information of the PV panel on the manufacturer's website via internet. Since micro-inverters have maximum power point tracking (MPPT) feature, they transfer the electrical power obtained from the panel to the network at the maximum level. Current, voltage and power information received from the panels can be transferred via micro-inverters to the digital environment and the data can be monitored instantly on the internet 7-24 and these data can be transferred elsewhere. Besides, with remote access, the energy values generated by the panels can be monitored daily, weekly and monthly. It is possible to follow the data graphically, as well as to report for production. The energy transferred to the grid on a daily, weekly and monthly basis can be seen in Figure 5.



a- Daily Electricity Generation

b- Weekly Electricity Generation



c- Monthly Electricity Generation

Figure 5: Energy provided by inverters to the grid

PV Module temperature sensors are equipped with an adhesive cover and weatherproof cable. These sensors are placed behind the PV module and enabled the measurement of the module surface temperature. The temperature information from the ambient temperature sensor and the module surface temperature sensor, and the information of radiation obtained by the inclined surface and the horizontal surface from the radiation pyranometer are transferred to the computer via a data acquisition card as shown in Figure 6.

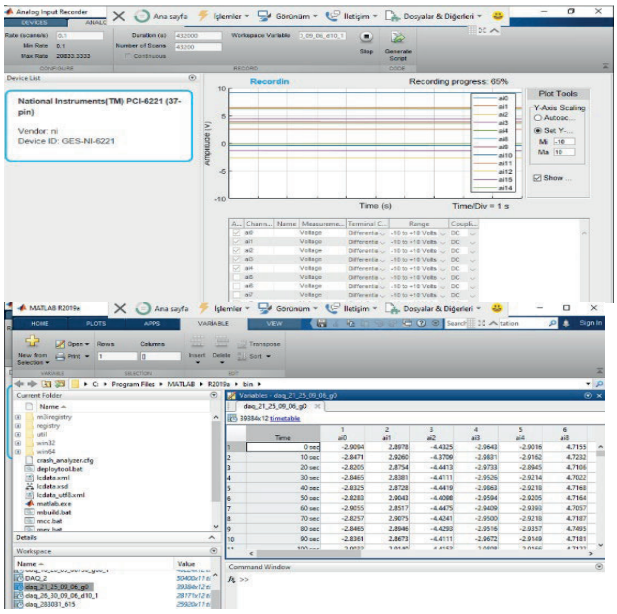


Figure 6: Radiance and Temperature data

III. Conclusion

Instantaneous, daily and total energy generated can be monitored with the experimental setup to determine the optimum inclination angle. Radiation values of different locations around the world are measured by local meteorological stations. The measured values are the radiation values of the surface. The irradiance values of the position and tilt of the panels are obtained with the pyranometers on this test setup. The irradiance values of the position and tilt of the panels are obtained with the pyranometers on this test setup. Thanks to these radiation values, it will be possible to determine a more accurate optimum angle of inclination and to develop a suitable mathematical model.

References

[1] M Yakup and A Q Malik, "Optimum tilt angle and orientation for solar collector in Brunei Darussalam", Renewable Energy, 24(2), 223-234, October 2001

[2] W D Lubitz, "Effect of manual tilt adjustments on incident irradiance on fixed and tracking solar panels", Applied energy, 88(5), 1710-1719, May 2011

[3] J Kaldellis, K Kavadias and D Zafirakis, "Experimental validation of the optimum photovoltaic panels' tilt angle for remote consumers", Renewable energy, 46, 179-191, October 2012

[4] M Kacira, M Simsek , Y Babur,and S Demirkol, "Determining optimum tilt angles and orientations of photovoltaic panels in Sanliurfa, Turkey", Renewable energy, 29(8), 1265-1275, July 2004

[5] K Bakirci,"General models for optimum tilt angles of solar panels: Turkey case study", Renewable and Sustainable Energy Reviews, 16(8), 6149-6159, October 2012

[6] N İ Beyazit, H Bulut and Y Demirtaş , "Yatay düzleme gelen difüz ışınım modellerinin ölçüm sonuçlarına göre analizi", Polytechnic Journal, 20(3): 557-563, September 2017

[7] M Tiris and C Tiris, "Analysis of solar radiation data for Gebze, Turkey", Energy conversion and management, 38(2), 179-186, January 1997

[8] M Tiris and C Tiris, "Optimum collector slope and model evaluation: case study for Gebze, Turkey", Energy Conversion and Management, 39(3-4), 167-172, February–March 1998

[9] M Kacira, M Simsek, Y Babur and S Demirkol, "Determining optimum tilt angles and orientations of photovoltaic panels in Sanliurfa, Turkey" Renewable energy, 29(8), 1265-1275, July 2004

[10] S S Chandel, R K Aggarwal and A N Pandey, "New correlation to estimate global solar radiation on horizontal surfaces using sunshine hour and temperature data for Indian sites", Journal of Solar Energy Engineering, 127(3), 417-420, September 2005

[11] G Lewis, "Optimum tilt of a solar collector Solar and Wind Technology", Journal Article 4(3), 407-410, January 1987

[12] Q Zhao, P Wang and L Goel, "Optimal PV panel tilt angle based on solar radiation prediction In Probabilistic Methods Applied to Power Systems (PMAPS)", 2010 IEEE 11th International Conference on (pp 425-430) IEEE June 2010

[13] A Rouholamini, H Pourgharibshahi, R Fadaeinedjad and G Moschopoulos, "Optimal tilt angle determination of photovoltaic panels and comparing of their mathematical model predictions to experimental data in Kerman" In Electrical and Computer Engineering (CCECE), 2013 26th Annual IEEE Canadian Conference on (pp 1-4) IEEE May2013

Temporomandibular Joint Sound Analysis with LSTM Networks

U. TAŞKIRAN[1], S. F. TAŞKIRAN[2] and M. ÇUNKAŞ[3]

[1]Selcuk University, Konya/Turkey, utaskiran@selcuk.edu.tr
[2]Konya Technical University, Konya/Turkey, sftaskiran@ktun.edu.tr
[3]Selcuk University, Konya/Turkey, mcunkas@selcuk.edu.tr

***Abstract* - Temporomandibular Disorder (TMD) is the problems arising from or related to disorder of Temporomandibular Joint (TMJ) which is commonly known as jaw bone joint. It is a common problem among the population, more than two third of population have different levels of TMD. TMD is mostly diagnosed by listening TMJ sounds. The procedure is an easy and cheap method used by clinic dentists. In this study features extracted from Fourier transformation of TMJ sounds. After reducing the size of feature vector using PCA, vectors are applied to Long-Short Term Memory (LSTM) neural networks to be trained and tested. Over 95% training accuracy and around 65 to 70% testing accuracy is achieved.**

***Keywords* - TMJ, TMD, sound classification, Recurrent Neural Networks, LSTM networks**

I. INTRODUCTION

TEMPOROMANDIBULAR joint (TMJ) is the joint which allows the movement of the jaw. TMJ is the joint between the condyle of the mandibula known as lower jaw or jawbone and mandibular fossa of the temporal bone. TMJ is very special joint and is an articular disc between these two bones enables jaws to make many complex movements. Many degrees of pain and movement loss may happen in the joint and all these are generally classified as Temporomandibular Disorder (TMD). In Figure 1 one can see the localization of TMJ and graphical explanation of TMD.

TMD frequency is very high in common population and may be as high as 75%. Generally accepted frequency of who shows one or more symptoms is around 33% [1]. These numbers are even higher in Turkey according to Özan at all [2]. Symptoms generally appear as reflected pains which appear to be not emanating from TMJ. These reflected pains can be mistaken as face pains, outer ear pains, face muscle pains, migraine type aches, even strains in the eyes etc. Because of this reason, patients generally seek help outside of dentistry area although TMD is generally resulting from teeth deformations. Disorder can be more severe like losing jaw movement range, mandibula deviations and deflection, pain and sound during opening and closing jaw, even jaw locks in extreme cases [1].

Main TMD diagnostic is clinical examination. At some point in clinical examination, a dentist listens the sound of the TMJ while the patient performs jaw opening and closing by using a stethoscope. Because jaw sounds is one of the main diagnostic lead of the disorder many academic studies focused on sound classification. The importance of sound based diagnose arises from its simplicity. Moreover it is readily available and cheaper than using sending the patient for MR imaging which is both costly and may not be available in time. TMD associated sounds are usually classified as crepitation, clicking, popping, popping clicking etc by dental community.

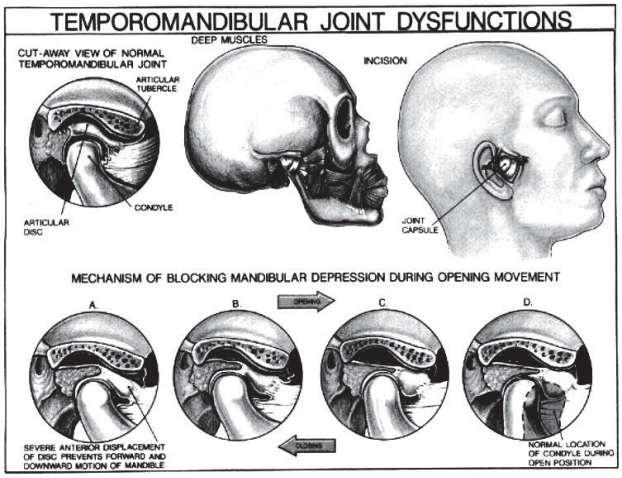


Figure 1: TMJ and TMD.

One of the early and known studies performed Wildmalm and his several different colleagues. In their study they first tried to classify the sounds like clicking, crepitation. They classify five different sounds by using Reduced Interference Distribution (RID) [3], [4]. Later studies by Wildmalm and his colleagues focused on amplitude wave spectrum of the TMJ sounds, classification based on pattern recognition [5], [6], [7].

Many different studies were focused on TMJ sound classification. These are includes the studies of Taskiran and his colleagues who studied classification of TMJ sounds based on frequency and statistical features based feature extraction

and ANN classification [8], [9]. In another study focused on Deep Learning Network, Taskiran is used a Convolutional Neural Network based image classification technique. Data is first preprocessed and turn into time-frequency images. Then the images are classified by using a image classification with very good results [10]. Akan at al. [11], [12] and Ghodsi at al. [13] published papers on classification of TMJ sounds using different classification techniques including spectrum based techniques, time-frequency analysis, discrete evolutionary transform. Similar techniques of sound classification are not just only applied for TMJ but also can be applied for knee joints to diagnose knee joints by Kim at al. [14]. Different objective studies of the same area are also published aiming the differentiating sound from left and right TMJ based kurtosis and click sounds [15], [16].

Electrocardiogram (ECG) and Electroencephalogram (EEG) classification was very active research area and still actively studied by many different academicians. Recently many researchers focus on and are popularize the deep learning classification methods. Recurrent Neural Networks (RNN) and Convolutional Neural Networks (CNN) are used for time series analyses. Among researchers some have used LSTM based classification procedures for ECG and EEG. LSTM is short for Long Short Term Memory and actually is an improved version of RNN. Wang et al. have used LSTM networks to classify EEG in motor imagery tasks [17]. Similarly Nagabushanam at al. have also used LSTM and an improved neural network for EEG classification and compared the results with other frequently used techniques [18]. Similar classification study was also performed by Saadatnejad at al. on the subject of ECG classification for continuous monitoring of wearable devices based on LSTM networks [19].

In this study we tried to classify previously recorded TMJ sounds by using LSTM network based techniques. Sound based diagnostic is one of the cheap methods to use in many application areas especially in medicine. Not only in dentistry but also in many areas of medicine listening the body of the patient are commonly used for diagnosis. Unfortunately diagnosing by listening requires expert ears for correct diagnostic. In this point, helping the dentist with a diagnostic tool will be very useful. Sound data has been recorded by using special developed tool and software in supervision of experienced dentists in real clinical environment. The used data is the same data of reference [8] and reference [10].

The process of data collection is thoroughly narrated and explained in aforementioned references. After preprocessing of data is performed, LSTM tools of MATLAB are used for classification. Fairly acceptable results are obtained and compared to previously obtained results. Further classification methods are planned to be tested in future and hopefully more robust real time diagnostic machine and software is expected to be developed.

II. Material and Method

As previously mentioned data set are from the previous works of the authors. Detailed collection procedure and designed tools and software explained in reference [10]. Summary of the procedure is given below. With the supervision of expert dentist of orthodontia, 5 seconds of TMJ sound data is recorded while the patients and test subjects opening and closing their mouths in real clinical environment. Sound data is sampled by 51200 Hz sample rate with 24 bits of resolution. Every sound data is recorded as 3 different sets of right TMJ sound, left TMJ sound and environment sound. Every sound set has total of 3 different, 24 bit 256000 sample data. Also every recorded data has its own data and time stamps and their respective tags whether if record belongs to a TMD patient or not.

When spectrums of the sound data are examined it can be seen that there were almost no high frequency components. Accordingly 51.2 KHz sample rate data is filtered through with BPF having pass frequency limits of 100 Hz to 10 KHz on 51.2 KHz. Low frequency components are sacrificed because of 50 Hz power line noise. Following the filtering all data is sampled down to 20 KHz. Thus procedure reduces the data size tremendously without sacrificing the content because there is no data above between 10 KHz to 25.6 KHz after band pass filtering. In the study both right and left TMJ sounds but not the environment sound included in classification process.

There are total of 150 data sets. 96 sets of the data belong to unhealthy patients and 54 sets of the data belong to healthy test subjects.

After the preprocessing step vector size (100000 samples) of each data is reduced to 2500 using Principal Component Analysis (PCA) for more convenient calculation. Principal Component Analysis is a dimensionality reduction method mostly used for compressing images or reducing size of large datasets. It reduces data dimension by projecting each data point onto the first few principal components while preserving as much of the data's variation as possible. Principal components are eigenvectors got from covariance matrix of data. MATLAB’s built in PCA function is used realize this process. Data vectors are reshaped to 50x2000 matrices to obtain principal component coefficients and reshaped back to 1x2500 vector after process. Then each data is converted to a MATLAB cell array to input them network.

Training and test data randomly selected from both healthy and unhealthy data. Then these healthy and unhealthy data are combined together and shuffled once more. A LSTM network is created with a sequence input layer, 10 bilstm layer, 2 fully connected layer, one sofmax layer and one classification layer. Network training has learning rate of 0.01 and training parameters are 5 epochs, 32 mini batch and 500 iterations. To validate the results script execution repeated 100 times correct classification and test rate are saved. Lastly, means of training classification and test classification results are calculated.

III. RESULTS

The MATLAB script is run for training and then data is tested. After properly adjusting the parameters to training the 100 run training and testing is done at the same time to confirm the results. At every script execution correct classification and test rate are saved. Finally, average of training classification and test classification results are computed. The obtained results are given in the Figure 1, Figure 2, and Figure 3. Note that H here is unhealthy and S is healthy.

According to the results training is highly successful to classify the data. Classification is 97% correct. 76 unhealthy data is used and training net correctly classifies 74 of them unhealthy and 2 of them healthy. 44 healthy data is used and training net correctly classifies 43 of them healthy and 1 of them unhealthy. Unfortunately, High success of training process is not reflected in test data. When the test data is presented to the trained net, success rate of the net is 70%. This, unfortunately, is because of limited number of data available for both training and test.

Some of the previous results are given in the Table 1 for comparison. LSTM seems to less successful method among previously applied methods of Frequency based Feature Extraction ANN, Statistics based Feature Extraction ANN, Deep Learning Image Based CNN.

Table 1: Comparison of previous results with LSTM

Method	Frequency based Feature Extraction ANN	Statistics based Feature Extraction ANN	Deep Learning Image Based CNN	Deep Learning Time Series LSTM
Simplified Success Rate	78%	89%	94%	70%

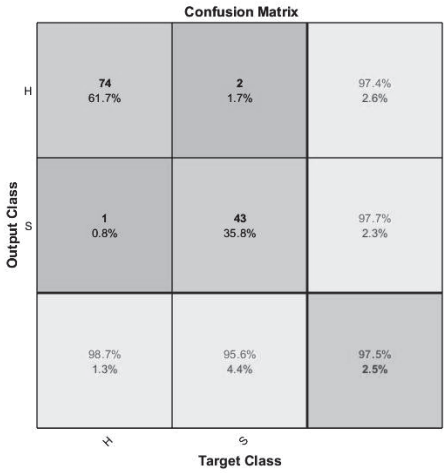


Figure 2: 100 run training confusion matrix.

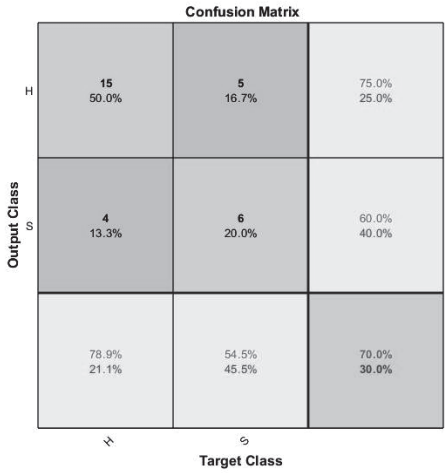


Figure 3: 100 run test confusion matrix.

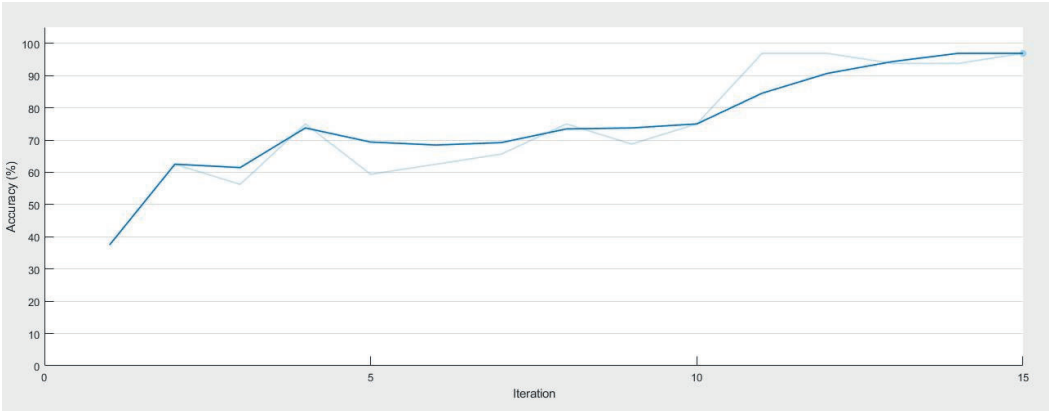


Figure 3: 100 run test accuracy graph.

IV. Conclusion

Aim of this study to explore the TMJ sound data classification capacity of LTSM networks which are one of the highly popularized deep learning tools. For this purpose previously obtained data sets of the author are used. Performance of the LSTM network tried to evaluate.

LSTM toolbox of MATLAB is used as software development base. A few scripts is used to preprocess the data and then by using preprocessed data LSTM network are trained and tested. According to results LSTM network trained very successfully but does not performed as good as expected. As seen from the results, although the accuracy of the LSTM training results is over 95%, test accuracy of 70%.

These results show that at LSTM training phase network memorizes the presented training data. This is generally consequence of insufficient number of data or data similarity of the sets. For more accurate results, gathering of more data sets of different test subject is required.

Being an ongoing research and development process, in future we are planning to obtain more data from healthy and unhealthy subjects. Also some data preprocessing will be applied to data to increase the training and test success rate. Improvement on LSTM algorithm is also planned.

Future study will also include sounds of different type of disorders. Identification of unhealthy joint by using sound data is also proposed.

References

[1] *Temporomandibular Disorders, Guidelines for Classification, Assessment and Management,* The American Academy of Orofacial Pain, Edited by C McNeil. Quintessence Books, Illinois 1993

[2] F Özan, S Polat, İ Kara, D Küçük, H B Polat, "Prevalence Study of Signs and Symptoms of Temporomandibular Disorder in a Turkish Population", *The Journal of Contemporary Dental Practice*, vol 8, no 4, May 2007

[3] S E Widmalm, W J Williams, B S Adams "The Waveforms of Temporomandibular Joint Sound Clicking and Crepitation " *Journal of Oral rehabilitation* 23; 44-49, 1996

[4] S E Widmalm, W J Williams, R L Christiansen, S M Gunn, D K Park "Classification of Temporomandibular joint sounds based upon their reduced interference distribution " *Journal of Oral rehabilitation* 23; 35-43, 1996

[5] Sano T , Widmalm S E , Westesson P L , Takahashi K , Yoshida H , Michi K , & Okanao T , "Amplitude and frequency spectrum of Temporomandibular joint sounds from subjects with and without other signs/symptoms of Temporomandibular disorders", *Journal of Oral rehabilitation* 26; 145-150, 1999

[6] Djurdjanovic D , Widmalm S E , William J W , Christopher K H Koh, and Kok Pin Yang "Computerized Classification of Temporomandibular Joint Sounds", *IEEE Transactions on Biomedical Engineering* Vol 47, No 8, August 2000

[7] Zheng C , Widmalm S E , Williams W J , "New Time-Frequency Analyses of EMG and TMJ sound signals " *IEEE Engineering in Medicine & Biology Society* 11th Annual International Conference 1999

[8] U Taşkıran, S Herdem, M Çunkaş, F Aykent, T Y Savaş, H E Koçer, "A Sample Reduction and ANN Based Classification of Temporomandibular Joint Sounds" International Research Journal of Electronics and Computer Engineering, Vol 1 No 3 (2015) (A pressed version of paper presented at Third International Conference on Science, Technology, Engineering and Management (3rd ICSTEM 2015)

[9] Taşkıran U , Taşkıran S F , Çunkaş M , 2018, Statistical Feature Extraction and ANN Based Classification of Temporamandibular Joint Sounds International Conference on Engineering Technologies (ICENTE'18) Oct, 26-28 2018 Konya, TURKEY

[10] Taşkıran U , "Temporomandibular Eklem Bozukluklarının Belirlenmesinde Sinyal İşleme ve Yapay Zeka Tekniklerinin Kullanılması " Ph D Dissertation Dept of Electrical-Electronics Engineering, Graduate Education Institute, Konya Technical University, Konya, TURKEY, 2019

[11] Akan A , Başar Ünsal R , "Time-frequency analysis and classification of Temporomandibular sounds" *Journal of The Franklin Institute*, 337 (2000) 437-451, 2000

[12] Akan A , Ergin A , Yildirim M , Öztaş, E , "Analysis of Temporomandibular joint sounds in orthodontic patients" *Computers and Electrical Engineering* 32 (2006) 312-321, 2006

[13] Ghodsi M , Hassani H , Sanei S , Hicks Y , "The use of Noise Information for detection of Temporomandibular disorder" *Biomed. Signal Process. and Control* vol 4, 79-85, 2009

[14] Kim, K S , Seo J H , Kang J U & Song C G , "An enhanced algorithm for knee joint sound classification using feature extraction based on time-frequency analyses" *Computer Methods and Programs in Biomedicine* 94, 198-206, 2009

[15] Took C C , Sanei S , Chambers J , & Dunne S , "Underdetermined blind source separation of Temporomandibular joint sounds" *IEEE Transactions on Biomedical Engineering* vol 53, No 10 Oct 2006

[16] Took C C , Sanei S , Rickard S , Chambers J & Dunne S , "Fractional delay estimation for blind source separation and localization of Temporomandibular joint sounds" *IEEE Transactions on Biomedical Engineering.* Vol 55No 3, Mar 2008

[17] Wang P, Jiang A, Liu X, Shang J, Zhang L "LSTM-Based EEG Classification in Motor Imagery Tasks " *IEEE Trans Neural Syst Rehabil Eng.,* 2018 Nov; 26(11): 2086-2095

[18] Nagabushanam, P , G S Thomas and R Sankararajan (2019) "EEG signal classification using LSTM and improved neural network algorithms " *Soft Computing* 24 10 1007/s00500-019-04515-0

[19] S Saadatnejad, M Oveisi and M Hashemi, "LSTM-Based ECG Classification for Continuous Monitoring on Personal Wearable Devices," in *IEEE Journal of Biomedical and Health Informatics*, vol 24, no 2, pp 515-523, Feb 2020

Design and fabrication of amorphous silicon based bottom gate thin film transistors and performance dependence on post annealing temperature

T.TUT[1], S. MEMİŞ[2], A. O. KODOLBAŞ[2], O. YILMAZ[2], M. AHMADLOO[4], A. KÖKSOY[2], E. MERT[3], F. KOCAMAN[3], and S. DABAK[2]

[1] Abdullah Gül University, Kayseri/Turkey, turgut.tut@agu.edu.tr
[2]TUBITAK Marmara Research Center, Gebze/Kocaeli,/Turkey, sema memis@tubitak.gov.tr
[3]Gebze Technical University, Gebze/Kocaeli/Turkey, elcinmert@gtu.edu.tr
[4]Bogazici University, Bebek/Istanbul/Turkey, ahmadloo1986@yahoo.com

***Abstract* - In this work, we report development of bottom gate AMOLED compatible thin film transistor micro fabrication processes and thin film material optimization process for device grade hydrogenated amorphous silicon, SiN_x and ALD Al_2O_3. Process problems were also pointed while we developed microfabrication. Post annealing remarkably improves the TFT performance up to 250°C beyond which it drops off. Observations were explained on the basis of formation of Schottky barrier material and diffusion of hydrogen out of amorphous silicon thin film.**

***Keywords* – TFT, thin films, microfabrication, display, transistors**

I. Introduction

Currently, large area electronics is one of the most important fields of research around the world. This technology has been used in different electronic devices, such as solar cells and Active Matrix Liquid Crystal Displays (AMLCDs) among others [1, 2]. The AMLCDs have been widely used in laptops, digital cameras and cellular phones. It is well known that the a-Si:H thin-film transistor (TFT) is the dominant pixel switching device in those displays [1–3].

Hydrogenated amorphous silicon (a-Si:H) has been very dominant recently for large area electronics for the following main reasons: Amorphous silicon has all the useful properties of a semiconductor including doping photoconductivity, field effect modulation of conductivity, easy junction formation, extensive knowledge attained about the chemistry of crystal silicon micro-fabrication processes over the years. The PECVD (Plasma Enhanced Chemical Vapor Deposition) process allows for low cost and fairly uniform growth of materials over large area substrates. PECVD process can be used to deposit, dielectrics (e.g. SiN_x, SiO_2 etc.), passivation layers, undoped and doped semiconductors with different energy band gaps (a-Si:H etc.). Thus, amorphous silicon is widely used in active matrix liquid crystal [1] displays and organic light emitting diode displays, scanners [2], position sensors [3], medical imagers [4], solar cells [5], biochemical sensors, stress and strain sensors.

In this letter, we report the deposition of high quality a-Si:H, SiN_x thin films by modified PECVD system, micro-fabrication development of a-Si:H based high performance thin film transistors for the use of AMOLED displays. State of the art quality a-Si:H thin films were grown and optimized the micro-fabrication processes such as annealing temperature and the etching processes etc. We also report the annealing temperature effect on the electrical characteristics of TFTs.

II. Experimental

Bottom gate TFT micro fabrication steps were as follows; As a first step, 180nm thick ITO (Indium Tin Oxide) coated 2 inch square Eagle Glass is cleaned using first the substrate is placed in Aceton containing beaker and this is also inserted in a ultrasonic bath at 60°C. Then under the optical microscope the substrate surface is inspected. If the surface is still contaminated, special cleaning procedure is applied. After final inspection, the substrate is taken into micro fabrication processes.

A. Gate dielectric and a-Si:H deposition

Hydrogenated amorphous silicon (a-Si:H or a-Si) TFT's were fabricated with standard bottom channel passivated structure on 2" X 2" Corning Eagle glass substrate. SiN_x and a-Si layer were deposited in multi-chamber ultrahigh vacuum, plasma enhanced chemical vapor deposition system (UHV-PECVD) at 13.56 MHz. Prior to any deposition, allocated chamber was evacuated down to a base pressure of the order of 10-9 Torr. About 300 nm thick a-Si and 150 nm thick n+-a-Si were sequentially deposited at a substrate temperature, chamber

pressure, RF power density, pure Silane flow rate of 225ºC, 0.6 Torr, 16.4 mW/cm^2, 10 sccm, respectively. Device-quality intrinsic film properties were assured using SEM, photo-, dark conductivity, FTIR and optical transmission and reflection measurements [Tanaka, K., and Matsuda A., Materials Science Reports 2 (1987) 139]. Optical Emission Spectrum was used to monitor plasma during deposition suggest that monomeric species SiH is dominated, has low intensity and constant through the deposition of the intrinsic a-Si film (Figure 1). For the deposition of n+-a-Si films, 4 sccm 3% PH_3 in H_2 were added to plasma. As gate dielectric 200-250 nm SiN_x and 35-50 nm Al_2O_3 were successively deposited. HV PECVD deposition conditions of SiN_x were as follows; substrate temperature: 225ºC, chamber pressure: 0.6 Torr, rf power density: 16.4 mW/cm^2, Silane flow rate: 10 sccm, Methane flow rate: 70 sccm. Deposition conditions of the passivation SiN_x were same as above with 50 nm thickness. Al_2O_3 films were deposited in Atomic Layer (ALD) deposition system using trimethylaluminum (TMA; ($Al(CH_3)_3$)) and water (H_2O) at substrate temperature of 150ºC. The deposition pressure was maintained at 600 mTorr. N_2 was used both as the precursor carrying and purging gas.

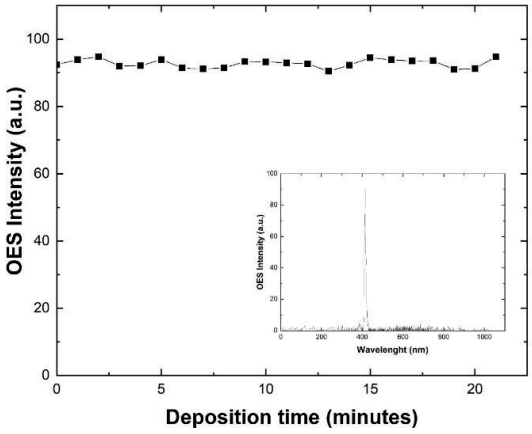


Figure 1: Evaluation of [SiH] peak at 413 7 nm with deposition time during the deposition of intrinsic amorphous silicon Insert shows that Optical Emission Spectrum (OES) of the same film

Capacitors and transistors were electrically shorted due to pinholes in the amorphous SiNx dielectric layer. We also noticed that there are spikes after the metallization process on the metal layer. We detected these spikes with profiler and optical microscope. Therefore, instead of just one SiNx dielectric layer as the gate dielectric material, we used 35-50 nm thick Al_2O_3 layer which was deposited with ALD (Atomic layer deposition). This Al_2O_3 thin film is used as a patching layer to cover the pinholes in the SiNx layer. This process leads to a great improvement in the performance of the TFTs and capacitors. This was the main problem we had to overcome to complete the microfabrication of TFTs successfully.

B. Microfabrication of TFT devices

In this section, we present a more detailed process flow for the fabrication of bottom gate thin film transistors. We design the photomasks such that it is possible to fabricate TFT, resistor, capacitor, and OLED monolithically in the same fabrication process. Therefore the TFTs, we have developed, can directly be used in AMOLED technology. However, in this paper we only report the TFT performance results. As a step 1, metal markers were deposited by Cr sputter with 100-200nm Cr thickness and then AZ1505 photoresist was used to transfer the marker patterns on the substrate. MJB4 SUSS mask alligner system was used for photolithography processes. Then after hard bake, Cr was patterned with Cr etcher solution. In the second step, since the etching recipe using wet chemistry is quite hard for ITO, BCl_3 based recipe is used to etch and pattern the ITO in order to create pads for OLEDs. We have to use different mask for this reason to etch a-Si:H and SiNx layers separately. In the third step, 200nm Cr is sputtered on the 2 inch square samples and then applied photolithography. Gate pads are patterned using Cr etcher solution under heat at 40ºC for 2 minutes.

After glass cleaning and drying process, 250nm SiN_x , 35nm Al_2O_3, 120nm undoped a-Si:H and 200nm n-type doped a-Si:H layers were deposited. SiN_x, and a-Si:H layers were deposited in a PECVD system in separate vacuum chambers . Al_2O_3 layer is deposited using ALD system. (The thicknesses were calculated by the test samples on which deposition processed is applied).

In the 4th step, AZ 5214E photoresist was used in photolithography in image reversal mode. By this way, markers and the substrate surface can be seen easily and the alignment process was facilitated. Since in this process, all the amorphous silicon layers and SiN_x were etched down to the glass layer by RIE etching, thick resist was used with approximately 1.5micron thickness. After image reversal photolithography, post bake at 110 ºC for 15 minutes was applied to prepare for the dry etching process. Using RIE with the CF_4/H_2 and BCl_3 based etching recipes, dielectric and amorphous silicon layers were etched down to the glass substrate except the transistor region. After etching process, AZ remover at 60 ºC for 1 hour is used to remove the remaining photoresist on the substrate. If the hard bake was done too much by using high temperature and for long time duration, then the resist removal would be problematic and some residues can stick to the substrate surface and lead to contamination. Amorphous silicon is also fragile material and therefore ultrasonic process cannot be used too long like 30 minutes and more at 120 ºC for stripping.

In order to retain the Al_2O_3 and SiN_x dielectric layers for the capacitor devices, we used two mesa masks.

In the 5th step, 150 nm thick SiN_x layer was deposited with UV PECVD system. Photolithography was also done with AZ1505 photoresist, the resist was baked at 100 ºC for 4 minutes before RIE etching process. The same CF_4/H_2 based RIE dry etching recipe was used to open the vias on SiN_x layer. To facilitate via formations, the vias over the source and drain contact regions were chosen to be between 10-20 micron widths. Vias were dry etched down to more than at least 120nm to make sure that all the SiN_x layer is etched. After the RIE

etching, AZ 100 remover was used again to remove the resist residuals on the surface at 60 ºC for 30 minutes time duration. Ultrasonic process was also applied as a last step to remove the residual resist for 15 seconds. The SiN_x deposited in this step is also used as the dielectric for the capacitors. The vias were open for source and drain contacts.

In the 6th step, we also used a separate via mask for OLED ITO pads and OLED pins at the edges of the sample. Since their areas are relatively larger than the openings for source and drain metal contacts, the dry etch rates will be larger for pins and OLED ITO pads.

In the 7th step, For the source and drain metal contacts, 150-200nm thick Cr metal was deposited by sputter on the whole sample. The important point in this deposition is that the side walls of the mesa structure should be covered with the thick enough Cr for the source and drain metal contacts otherwise there would be electrically open contacts. Using 5214 photoresist in image reversal mode, metal patterns were transferred onto the substrate. Cr wet etcher was used at 40 ºC for about 2 minutes to etch down the Cr metal. The resulting TFT device is depicted in Figure 5.

In the 8th step, AZ 5214E photoresist were used for dry etch mask. In this RIE etching process, SiNx and n+ a-Si:H layers has to be etched to reach the n channel layers and separate the source and drain n+ layers only and all other areas are protected with photoresist. In Figure 6 the TFT under microprobes has 12micron gate length.

In the last step, thick 600nm SiNx layer is deposited using HV PECVD system. Then using AZ5214E photoresist masks open the large vias (100-200 micron wide) for the Pins and ITO OLED pads. The schematical cross sectional view for the completed TFTs isshown in Figure 2 and the optical microscope image of a complete device under microprobes is shown in Figure 3.

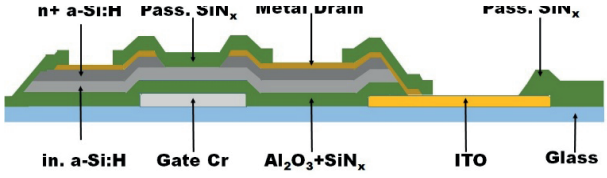


Figure 2: Complete TFT device schematics with ITO integrated

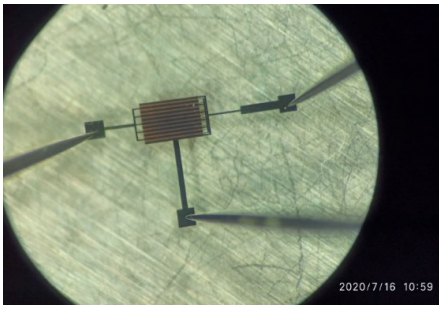


Figure 3: Optical microscope image of complete TFT under test, W=255µm L=12µm F=10

C. I-V Measurements

We used four probe station to connect the device under test and 4200-SCS Keithley Semiconductor Parametric Analyzer as SMU unit. The device under test is probed using microprobes. Microprobe needles have 5 micron diameter. Measurements are conducted in a dark room in order to prevent extra photocurrent to affect our electrical measurement. The field effect mobility in the saturated region was calculated from the Ids-Vds plots. For the annealing optimization process, we first annealed the TFT chips under vacuum at 150ºC for 30 minutes. Then conducted the I-V measurements. We continued to conduct the I-V measurements by increasing temperatures. We tried 200 ºC and 250 ºC for the optimization. When we used 200 ºC, we get the optimum result in terms of drain current. After 250 ºC generally the devices's performance get worse in terms of drain current and stability. Generally from the transistors annealed at 200 ºC, approximately we measured electrical parameters were Vth=0.9V, Ids(on)/Ids(off)=10^4-10^5. It has saturation mobility of about 0.08 cm^2/V.s. [Figure 4]. First reason for that actually comes from the fact that, the E-field distribution is better around the channel region, when transistors have multiple gate finger structures connected in parallel. Of course as the device area increases, the device switching speed decreases. However, for display technology, speed is not so critical, low threshold voltage, on/off ratio of drain current, and leakage current are more important. Another reason is that as the anneal temperature increases the ohmic contacts between the source and drain metals and the n+ a-Si:H layer get better and this results in an increase in drain current under the same Vds voltage. However, when the anneal temperature is increased further, the contacts get worse and the stability of the device deteriorates. This can be explained as there are two main factors affecting the ohmic contacts of the interface between chromium metal and the Silicon surface. The annealing process during the growth of the thin films inside the deposition chamber and annealing process after the devices are complete have different effects on the ohmic contacts. High temperature process during the growth of a-Si:H and SiN_x films results in high quality a-Si:H and SiN_x films and TFTs with high mobility, low leakage, and high device stability[6].As a metal is deposited on silicon, a silicide-silicon interface is usually formed between the metal and the silicon interface. Si forms silicides with most metals even at temperatures below 200 ºC[7,8]. Thus, the silicide-silicon interface lead to a rectifying barrier or an ohmic contact to silicon films. Common ohmic contact materials to silicon are chromium, molybdenum, and tungsten [9,10,11]. Cr/a-Si:H interlayer is amorphous and its composition is mostly $CrSi_2$ at about 300ºC. As the anneal temperature of the vacuum bake is increased until 250 ºC, the mobility and the resulting Ids increases however as the anneal temperature is increased further above 250º C, the Shottky potential barrier and the diffusion of hydrogen content in the amorphous silicon results in the deterioration of ohmic contacts and also a decrease in the mobility and drain current. [12]

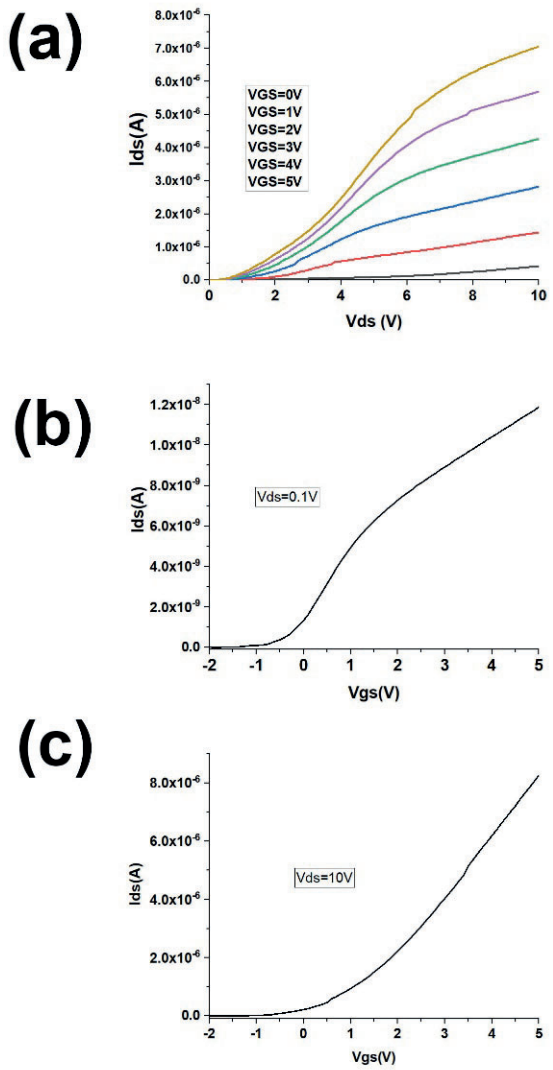


Figure 4: TFTs I-V characteristics for W=1000um, L=12um, F=10um

III. Conclusion

In this work, deposition of high quality undoped, n-type doped hydrogenated amorphous silicon, and SiN_x thin films growth were developed successfully. Using these thin films, amorphous silicon based TFTs with bottom gate architecture with 5um gate length was fabricated. This are one of the smallest amorphous silicon based TFTs fabricated in Turkey. The fabrication processes especially reported in detail to demonstrate the critical microfabrication process problems and the solutions that were developed by our research group. Photolithography, wet etching, dry etching, RF sputtering processes were explained in detail. The completed TFT devices were characterized electrically and high performance was achieved in terms of drain current, threshold voltage and Id(on)/Id(off) current. The annealing temperature effect was also investigated. As the temperature was increased from 100 ºC to 200 ºC, the Ids current increased for the TFT devices but as the temperature was higher than 250 ºC, the device performance degraded and the devices become unstable. As future work, we will fabricate bottom gate TFTs with OLED, TFT pixels. We expect better electrical performance for bottom gate TFTs when we apply channel passivation process. We will try to change the E-field distribution so that drain current will be higher at a given Vds voltage.

Acknowledgment

This work is partially supported by Scientific and Technological Council of Turkey under the project coordination number 113G035. Authors would like to thank S. P. Mucur for ALD coatings who works in the OLED fabrication team in the project, to Cihat Taşaltın for I-V measurement who works at Tubitak MAM BBB, to Şafak Gök for photolithography, who works at GTU (Gebze Technical University), used in this work.

References

[1] P G LeComber , W E Spear and A Gaith, "Amorphous-silicon field-effect device and possible application," Electronic Letters, vol 15, pp 179-181, 1979

[2] R C Chittick, J H Alexander and H F Sterling, "The Preparation and Properties of Amorphous Silicon," Journal of Electrochemical society, vol 116, pp 77-89, 1969

[3] P G LeComber and W E Spear, "Electronic transport in amorphous silicon Films," Phys Rev Lett , vol 25, pp 509-511, 1970

[4] W E Spear, R J Loveland and Al-Sharbaty, "The temperature dependence f photoconductivity in a-Si," Journal of Noncrystalline Solids, vol 15, 1973

[5] R A Street, Hydrogenated Amorphous Silicon Cambridge University Press, 1991

[6] K Long, A Z Kattamis, I-C Chang, H Gleskova, S Wagner, and J C Sturm, "Reliability of Active-Matrix Organic Light-Emitting-Diode Arrays With Amorphous Silicon Thin-Film Transistor Backplanes on Clear Plastic," IEEE Electron Devices Letters, vol 27, No 2, pp111-113, 2006

[7] 1-K N Tu, "Selective growth of metal-rich silicide of near-noble metals," Appl Phys Lett 27:221, 1975

[8] J W Mayer and K N Tu, "Analysis of thin-film structures with nuclear backscattering and x-ray diffraction," J Vac Sci Technol , 11(1):86, 1974

[9] M Akiyama, Y Ikeda, M Ikeda, and K Suzuki, Society for Information Display Digest, 24:887, 1993

[10] J W Mayer, and S S Lau, Electronic Materials Science for Integrated Circuits in Si and GaAs, D Johnstone Ed , Chap 11, Macmillan Publishing Company, 1990

[11] R Turan and N Akman, "Schottky barrier height of $CrSi_2$-Si junctions," Semicond Sci Technol , 8:1999, (1993)

[12] S J Park, Y T Sul, B S Bae, Y S Kim, J W Choi, J Jang, J-H Choi, J C Goh, and K Chung, "Stability of a-Si:H TFT after high temperature annealing," Proc of ASID '06, 8-12 Oct, New Delhi

An Electronically Tunable Memcapacitor Based on Floating Current Source

M. YILDIRIM[1]

[1] The Scientific and Technological Research Council of Turkey (TUBITAK), Ankara/Turkey, melih.yildirim@tubitak.gov.tr

Abstract **- In this work, an electronically tunable memcapacitor emulator circuit has been proposed. The memcapacitor consists of one Floating Current Source (FCS), two capacitors, one MOS transistor and one dependent voltage source. In emulator circuit, a MOS transistor is utilized in order to make the memcapacitor electronically tunable. Different values of voltage are applied to the gate terminal of the MOS transistor to show the property of electronically tunability of the memcapacitor emulator. The FCS circuit includes only four MOS transistors. In addition, the suggested memcapacitor emulator circuit does not require any mutator for the purpose of converting the emulator structure from memristor to memcapacitor. The characteristics of the electronically tunable memcapacitor are presented using LTspice simulation software and 0.13µm IBM CMOS technology parameters are used in circuit design.**

Keywords **– memcapacitor, electronically tunable, floating current source (FCS).**

I. INTRODUCTION

In 1971, Leon Chua postulated the fourth passive electronic component called memristor, which is abbreviation of memory and resistor, differently from other passive elements which are resistor, capacitor and inductor [1]. The memristor provides a constitutive relationship between charge and flux, whereas the resistor provides a constitutive relationship between voltage and current, the capacitor provides a constitutive relationship between voltage and charge and the inductor provides a constitutive relationship between current and flux [2].

The first physical realization of memristor, which is composed of two titanium dioxide (TiO_2) thin film layers, was done by a research group in Hewlett-Packard (HP) lab in 2008 [3]. After the fabrication of the memristor, numerous studies on this passive element have been carried out in literature [4–13].

In addition to memristor, the other mem-elements which are memcapacitor and meminductor have also attracted the researchers' interest. Many studies with regard to memcapacitors [2,14–18] and meminductors [14,19–21] have been done. A study carried out by Babacan [14] has introduced an Operational Transconductance Amplifier (OTA) based memcapacitor. However, the memcapacitor emulator does not exhibit the property of electronically tunability.

In this study, the emulator circuit proposed in [14] has been modified and FCS based memcapacitor emulator circuit has been suggested. In literature, many studies [17,21] on memcapacitor and meminductor require mutator for transforming from memristor to memcapacitor and meminductor. The presented emulator in this study does not require any mutator in order to transform from memristor to memcapacitor. Madsar et al. [18] have suggested a FCS based memcapacitor which can not be adjusted electronically. In this study, differently from the memcapacitor emulators in [14,18], the proposed emulator has the ability of electronically tunable.

II. ELECTRONICALLY TUNABLE MEMCAPACITOR EMULATOR STRUCTURE

In this study, the memcapacitor emulator suggested in [14] has been modified in order to achieve the ability of electronically tunability. For this purpose, an additional MOS transistor is utilized in the design. In addition, the memcapacitor emulator in this study is based on FCS. As is known, FCS is a current mode structure. In the suggested emulator structure, FCS is preferred since current mode circuit structures offer the behaviour of very low power consumption. Figure 1 demonstrates FCS structure which includes only four MOS transistors. Using Kirchhoff's current law, Eqs. (1)-(3) are obtained from FCS circuit structure [22]. The aspect ratios of MOS transistors of FCS structure are given in Table 1.

$$I_{B1} - I_{B2} - \left(I_{OUT\_P} + I_{OUT\_N}\right) = 0 \quad (1)$$

$$I_{B1} = I_{B2} = I \quad (2)$$

$$I_{OUT\_P} = -I_{OUT\_N} \quad (3)$$

Table 1: Aspect ratios of MOS transistors and bias current values for FCS

Parameters	Values
W/L for M_1-M_2	160µm/0.13µm
W/L for M_3-M_4	110µm/0.13µm
I_{B1}=I_{B2}	100mA

The electronically tunable FCS based memcapacitor emulator is given in Figure 2. The emulator circuit has one FCS, two capacitors, a MOS transistor and a dependent voltage source. FCS structure used in emulator has one positive input, one negative input and three positive output terminals. In FCS structure, the capacitor C_1 is connected to the positive input terminal. The source terminal of M_5 transistor is connected to the negative input terminal of FCS. By applying V_G to the gate terminal of M_5 transistor, the proposed memcapacitor acquires

the property of electronically tunability. A dependent voltage source is connected to drain terminal of M_5 transistor. The voltage source depends on the multiplication of the voltage across the capacitor C_2 and its integral [14,16,18]. In FCS structure, one of the output terminals is connected to the positive input terminal. One of the output terminals of FCS is connected to the capacitor C_2 in order to provide the memory property of the memcapacitor. Another output terminal of FCS is used as the output of the memcapacitor emulator. Table 2 gives the circuit parameters belonging to FCS based memcapacitor emulator.

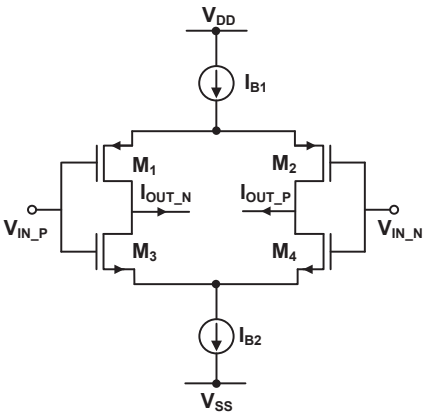


Figure 1: The simple MOS realization of FCS

Table 2: Circuit parameters of FCS based memcapacitor emulator

Parameters	Values
C_1	500nF
C_2	2nF
V_{DD}=-V_{SS}	1V
W/L for M_5	1µm/1µm

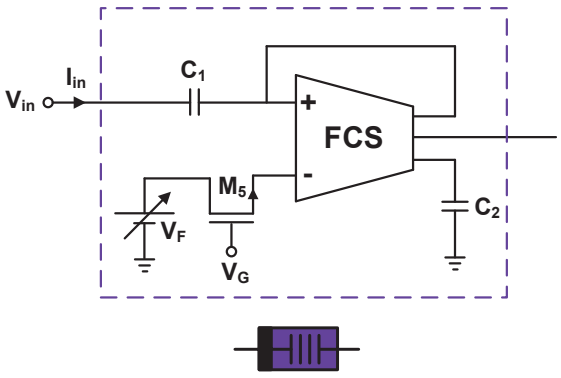


Figure 2: FCS based electronically tunable memcapacitor emulator

In FCS based electronically tunable memcapacitor emulator circuit, the input current I_{in} can be defined as

$$I_{in} = C_1 \frac{d(V_{in} - V_F - V_{SD})}{dt} \quad (4)$$

Eq. (4) can be expressed as given in Eq. (5).

$$\frac{q(t)}{C_1} + V_F + V_{SD} = V_{in} \quad (5)$$

$$V_F = V_{C2} \int V_{C2} = \frac{q(t)}{{C_2}^2} \int q(t)dt \quad (6)$$

where $q(t)$ is equal to $C_2 V_{C2}$ and it is variable part of suggested FCS based memcapacitor. The change in V_F provides the memcapacitive feature of the suggested memcapacitor emulator circuit.

III. Simulation Results

In this part study, DC transfer characteristics and transconductance of FCS structure are demonstrated. In addition, the graphics of input voltage-time, charge-time and voltage-charge relationships are given for suggested FCS based electronically tunable memcapacitor.

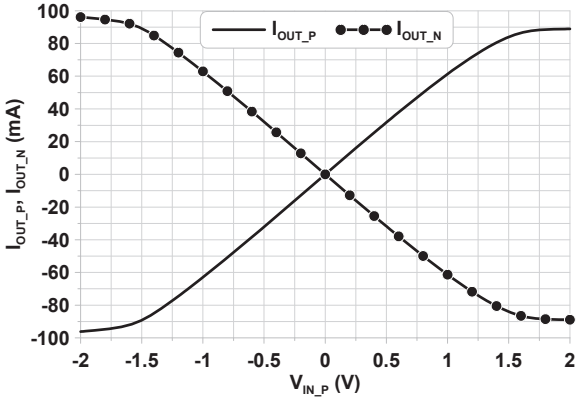


Figure 3: DC transfer characteristics of FCS structure

Figure 3 gives DC transfer characteristics of FCS. In order to plot the DC transfer characteristics, ±2V is applied to positive input. Positive and negative current outputs are obtained.

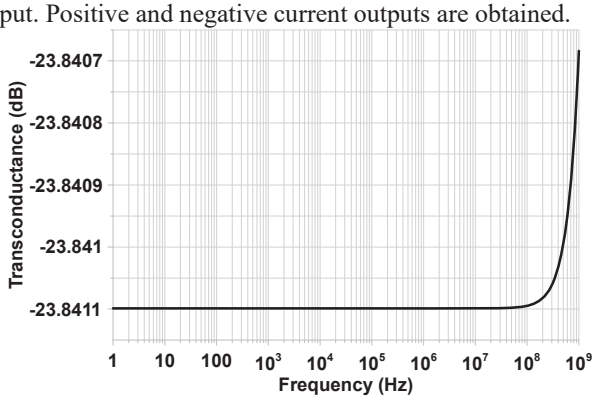


Figure 4: The transconductance value of FCS

Figure 4 shows the transconductance value of FCS circuit in frequency domain. The transconductance value of FCS is 64.26mS. It can be clearly seen that transconductance of FCS in Figure 4 agrees well with the slope generating the transconductance value in Figure 3.

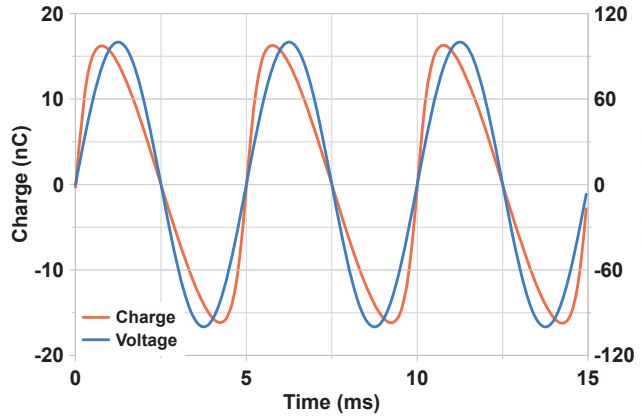


Figure 5: Charge and input voltage of proposed FCS based electronically tunable memcapacitor

Voltage-time and charge-time graphics of introduced memcapacitor are demonstrated in Figure 5. A sinusoidal 100mV input voltage with 200Hz frequency is applied to FCS in order to obtain voltage-time and charge-time graphics.

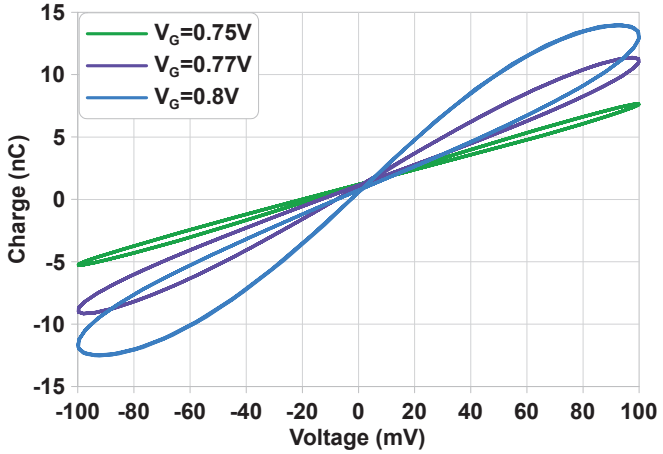


Figure 6: Voltage-charge relationship with various V_G values for f=200Hz and V_G=0.75V, V_G=0.77V, V_G=0.8V

In Figure 6, voltage-charge relationships with various V_G values are given applying sinusoidal 100mV input signals with 200Hz frequency. Different V_G values which are 0.75V, 0.77V and 0.8V are applied to show the behaviour of the electronically tunability.

In Figure 7, voltage-charge relationships with various frequencies are given applying 0.9V to V_G. Sinusoidal 100mV input signals with different frequencies such as 100Hz, 200Hz and 400Hz are applied to FCS in order to show the changing of voltage-charge curves depending on frequency value.

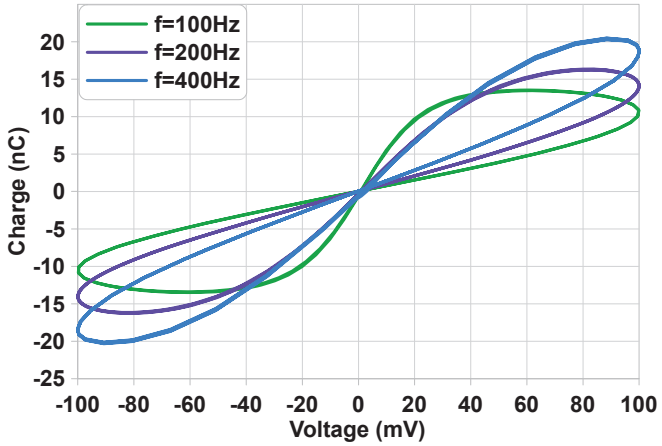


Figure 7: Voltage-charge relationship with various frequencies for V_G=0.9V and f=100Hz, f=200Hz, f=400Hz

IV. Conclusion

In this study, a FCS based memcapacitor emulator circuit is presented. The suggested emulator has the ability of electronically tunability thanks to applying V_G to the gate terminal of M_5 transistor. In literature, numerous memcapacitor circuits include memristor element. However, the introduced electronically tunable memcapacitor is generated without utilizing memristor element. In other words, the introduced emulator circuit does not require any mutator to transform from memristor to memcapacitor. Circuit simulations are carried out in order to show the charge-voltage relationship of the suggested memcapacitor. All simulation results agree well with the previous studies in literature.

References

[1] Chua L. Memristor-The missing circuit element. IEEE Trans Circuit Theory 1971;18:507–19. doi:10.1109/TCT.1971.1083337.

[2] Konal M, Kacar F. Electronically Tunable Memcapacitor Emulator Based on Operational Transconductance Amplifiers. J Circuits, Syst Comput 2020:2150082. doi:10.1142/s0218126621500821.

[3] Strukov DB, Snider GS, Stewart DR, Williams RS. The missing memristor found. Nature 2008;453:80.

[4] Yesil A, Babacan Y, Kacar F. An Electronically Controllable, Fully Floating Memristor Based on Active Elements: DO-OTA and DVCC. AEU-International J Electron Commun 2020:153315.

[5] Yildirim M, Babacan Y, Kacar F. Memristive retinomorphic grid architecture removing noise and preserving edge. AEU - Int J Electron Commun 2018;97:38–44. doi:10.1016/j.aeue.2018.10.001.

[6] Yesil A, Babacan Y, Kacar F. Electronically tunable memristor based on VDCC. AEU-International J Electron Commun 2019;107:282–90.

[7] Sözen H, Çam U. Electronically tunable memristor emulator circuit. Analog Integr Circuits Signal Process 2016;89:655–63. doi:10.1007/s10470-016-0785-2.

[8] Yildirim M. DNA Encoding for RGB Image Encryption with Memristor Based Neuron Model and Chaos Phenomenon. Microelectronics J 2020:104878.

[9] Babacan Y, Kaçar F, Gürkan K. A spiking and bursting neuron circuit based on memristor. Neurocomputing 2016;203:86–91. doi:10.1016/j.neucom.2016.03.060.

[10] Sánchez-López C, Mendoza-Lopez J, Carrasco-Aguilar MA, Muñiz-Montero C. A Floating Analog Memristor Emulator Circuit. IEEE Trans Circuits Syst II Express Briefs 2014;61:309–13.

[11] Petrović PB. Tunable flux-controlled floating memristor emulator circuits. IET Circuits, Devices Syst 2019;13:479–86.

[12] Yildirim M, Kacar F. Chaotic Circuit with OTA based Memristor on Image Cryptology. AEU-International J Electron Commun

2020:153490.
[13] Petrović PB. Floating incremental/decremental flux-controlled memristor emulator circuit based on single VDTA. Analog Integr Circuits Signal Process 2018;96:417–33.
[14] Babacan Y. An operational transconductance amplifier-based memcapacitor and meminductor. Istanbul Univ - J Electr Electron Eng 2018;18:36–8. doi:10.5152/iujeee.2018.1806.
[15] Yuan F, Li Y, Wang G, Dou G, Chen G. Complex dynamics in a memcapacitor-based circuit. Entropy 2019;21:188.
[16] Fouda ME, Radwan AG. Charge controlled memristor-less memcapacitor emulator. Electron Lett 2012;48:1454–5.
[17] Biolek D, Biolkova V. Mutator for transforming memristor into memcapacitor. Electron Lett 2010;46:1428–9. doi:10.1049/el.2010.2309.
[18] MADSAR M, BABACAN Y, ÇİÇEK K. FCS Based Memcapacitor Emulator Circuit. J Inst Sci Technol 2020:112–7. doi:10.21597/jist.670900.
[19] Konal M, Kacar F. Electronically tunable meminductor based on OTA. AEU-International J Electron Commun 2020;126:153391.
[20] Romero FJ, Escudero M, Medina-Garcia A, Morales DP, Rodriguez N. Meminductor Emulator Based on a Modified Antoniou's Gyrator Circuit. Electronics 2020;9:1407.
[21] Sah MP, Budhathoki RK, Yang C, Kim H. Mutator-based meminductor emulator for circuit applications. Circuits, Syst Signal Process 2014;33:2363–83.
[22] UGRANLI HG, YILDIRIM M, KAÇAR F. Design of low power DTMOS based FCS and its notch filter application for ECG signals. Int J Comput Exp Sci Eng n.d.;3:29–32.

Investigation of Image Compression Performance on Medical Images by Using Discrete Orthonormal Stockwell Transform and Singular Value Decomposition Methods

Gülhan USTABAŞ KAYA[1] and Tuğba Özge ONUR[1]

[1]Zonguldak Bulent Ecevit University, Zonguldak/Turkey, ozdincozge@gmail.com
[1]Zonguldak Bulent Ecevit University, Zonguldak/Turkey, gulhan.ustabas@beun.edu.tr

Abstract **-Image compression in medical images is a key step and the only applicable technique to prevent the consuming maximum storage, is to reduce the size of the data and utilize the maximum bandwidth of image transmission. Currently, there has been abundant research on image compression for medical images. In this study, Singular Value Decomposition (SVD) and Discrete Orthonormal Stockwell transform (DOST) techniques, which are recommended to be used frequently in the existing studies, are performed in order to obtain the best image compression without any loss of information for computed tomography (CT) scan slice of the brain and CT abdomen images and their performances are compared. Mean square error (MSE) and mean absolute error (MAE)- the used quality parameters- are computed to assess the performance of this study. The obtained results indicate that the usage of SVD for the compression of medical images is the best method in order to obtain the image closer to the original one.**

Keywords–**Image compression, medical image, singular value decomposition, discrete orthonormal Stockwell transform.**

I INTRODUCTION

In recent years advances, important studies have been carried out in the field of imaging diseases and diagnosis with digital technology. Medical imaging techniques such as digital radiography (DR), computed tomography (CT), computed radiography (CR) and magnetic resonance imaging (MRI) are used for accurate and reliable detection of diagnosis and treatment of diseases. However, large storage areas are required in these methods for image acquisition, storage and transmission of large volume images [1]. In addition, the spatial, temporal and spectral redundancies of the images cause problems such as the need for large storage space, slowing down speed of data and image transmissions. Image compression is a solution used to overcome these problems [2]. It decreases the capacity of storage space by maintaining the properties and quality of image. The aim of image compression is to keep the image quality by using the smallest possible number of bits. Because the uncompressed images or data require the transmission bandwidth and storage capacity.

To compress the images, two methods are applied. One of them is lossless compression, which is known as reversible compression. It does not compromise data loss. The other one is the lossy method called as nonreversible compression [3]. It is used in signal processing applications due to the fact that the providing the much greater compression with lossy compression [4]. Wavelet, higher-level JPEG, Discrete Cosine Transform, Singular Value Decomposition (SVD), Discrete Rajan Tranforms (DRT), Discrete Orthonormal Stockwell transform (DOST) are the examples of the used methods for lossy compression. In this context, there have been numerous lossy compression methods which are applied to medical images [3,5,6]. Cao applied the SVD for digital image processing application. Cao tested the face recognition and digital image compression by using SVD [7]. Wang, et.al investigated the image compression to provide a representation of multiresolution spatial- frequency and high image quality by using DOST [8]. Liu et. al discuss the current role of the image compression performances based on the standards with publicly available datasets in medical imaging applications [9]. Koff and Shulman overviewed groundwork for digital compression of medical image in radiology by using lossy image compression [10].

In this paper, the image compression for medical images- CT scan slice of the brain and CT abdomen image are used- is evaluated by using SVD and DOST techniques. We also compared the performances of these techniques with the calculation of mean square error (MSE) [11] and mean absolute error (MAE) [12] values between the original and reconstructed images.

The rest of the paper is organized as follows: The methodology is presented in Section 2. The SVD and DOST techniques are explained in detail. In section 3, we present the simulation results by examining the performances of DOST and SVD transforms on different medical images. The MSE and MAE values are calculated in this section to reinforce the comparison of the used methods. Finally, a brief discussion is given in Results and Discussion section called as Section 4.

II. METHODOLOGY

A. Singular Value Decomposition (SVD)

Since SVD is applicable to any matrix which is real and with $m \times n$ size, it is commonly used in linear algebra. A matrix with m rows, n colums and r rank can be reformed by three matrices as given in (1),

$$A = USV^{\mathrm{T}} \tag{1}$$

Here, S denotes diagonal matrix with $m \times n$ size and U and V are $m \times m$ and $n \times n$ orthogonal matrices, respectively. Equation (2) defines the U and V matrices,

$$\begin{aligned} U &= \left[u_1, u_2, ..., u_r, u_{r+1}, ..., u_m\right] \\ V &= \left[v_1, v_2, ..., v_r, v_{r+1}, ..., v_n\right] \end{aligned} \tag{2}$$

and the orthonormal set formed by u_i and v_i vectors are given in (3),

$$\begin{aligned} u_i^{\mathrm{T}} u_j &= \delta_{ij} = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \\ v_i^{\mathrm{T}} v_j &= \delta_{ij} = \begin{cases} 1, & \mathrm{i} = j \\ 0, & i \neq j \end{cases} \end{aligned} \tag{3}$$

The diagonals of the S matrix include the singular values as shown in (4),

$$S = \begin{vmatrix} \sigma_1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_r & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & \sigma_{r-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & \sigma_n \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{vmatrix} \tag{4}$$

where σ_i values indicate the singular values of A matrix for $i = 1,2,3,...,n$. The increase in the rank causes the decrease in the singular values of a matrix. Therefore, if the smaller singular values or the high ranks are eliminated, noise is reduced or the data is compressed. When SVD transform is used for an image, the great amount of the image information is in the first singular value. So, this enables the image to be represented by using only a few singular values. As a result, without any compression the image is obtained much close to the original one [13]. Equations given in (5)-(7) defines the image compression process by using SVD transform.

$$\begin{aligned} A = USV^{\mathrm{T}} &= \sum_{i=1}^{r} \sigma_i u_i v_i^{\mathrm{T}} \\ &= \sigma_1 u_1 v_1^{\mathrm{T}} + \sigma_2 u_2 v_2^{\mathrm{T}} + \cdots + \sigma_r u_r v_r^{\mathrm{T}} \end{aligned} \tag{5}$$

When performing the image compression, since the smallest singular values are dropped, summation is not applied up to the last singular values. Therefore, after a k number of term, these summation is truncated and the k rank of the matrix can be obtained as shown in (6),

$$A_k = \sigma_1 u_1 v_1^{\mathrm{T}} + \sigma_2 u_2 v_2^{\mathrm{T}} + \cdots + \sigma_k u_k v_k^{\mathrm{T}} \tag{6}$$

With this truncation, the total store space which is essential for A_k can be defined as,

$$k\left(m + n + 1\right) \tag{7}$$

The values of k will affect the formed image and thereby, the store space of that image. There have been various researches for determining the optimal k value in the literature, however, in general, the smaller k value than n, forms the resulting image closer to the original one [14].

B. Discrete Orthonormal Stockwell Transform (DOST)

The DOST is the orthonormal form of the Stockwell transform. The Fourier transform of the multiplication of a $h(t)$ signal by a Gaussian window calculates the Stockwell transform as defined in (8) [15].

$$S(\tau, f) = \int_{-\infty}^{\infty} h(t) \frac{|f|}{\sqrt{2\pi}} e^{-\frac{(\tau - t)^2 f^2}{2}} e^{-i2\pi f t} dt \tag{8}$$

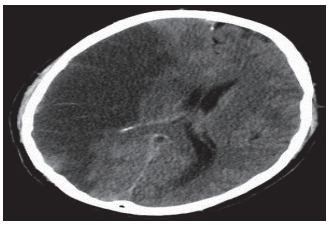
Here, f denotes the frequency, both t and τ indicate the time. However, if (8) is considered, it can be seen that ST has computational complexity [16]. Thus, the increase in the dimension makes the implementation of ST transform expensive. The DOST is a reduced form of the ST. Since longer periods are for lower frequencies and higher sampling rates are for high frequencies, the lower frequencies are subsampled by the DOST [17]. The basis vectors of DOST are given in (9),

$$S(k,t)_{[v,\beta,\tau]} = ie^{-i\pi\tau} \frac{e^{-i2\alpha\left(v - \frac{\beta}{2} - \frac{1}{2}\right)} - e^{-i2\alpha\left(v + \frac{\beta}{2} - \frac{1}{2}\right)}}{2\sqrt{\beta} sin\alpha} \tag{9}$$

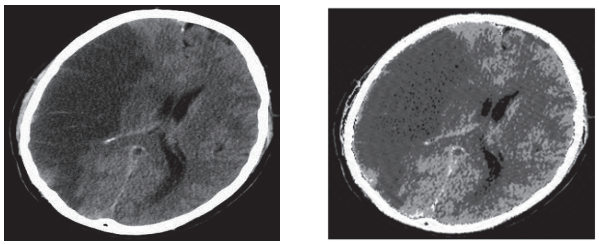
where N is the signal length, k=0,....,N-1, $\alpha=\pi(k/N-\tau/\beta)$ denotes the window center, v corresponds to the center of each frequency band and the bandwidth is represented with β . Therefore, in order to obtain an effective DOST transform, the parameters of β, v, τ have to be selected properly.

III. SIMULATIONS

In this paper, in order to examine the performances of DOST and SVD transforms on the medical images, two different medical images are used. A CT scan slice of the brain that has a right hemispheric cerebral infarct [18] and a CT abdomen image [19] are used for simulations performed in MATLAB. SVD transform and DOST techniques are applied on both CT brain and CT abdomen images. The original images and the results obtained by applying SVD transform and DOST technique to the original images are shown in Fig. 1 and 2. In Fig. 1, first row indicates the original CT scan slice of the brain and the obtained images by using SVD transform and DOST technique are shown in the left and right sides of the second row, respectively. It can be seen easily from Fig. 1 that the closest image to the original one is obtained with SVD transform.When DOST transform is used, the image distortion has occured and some details in the original image have lost in the obtained one.



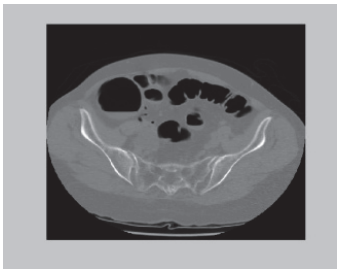
(a)



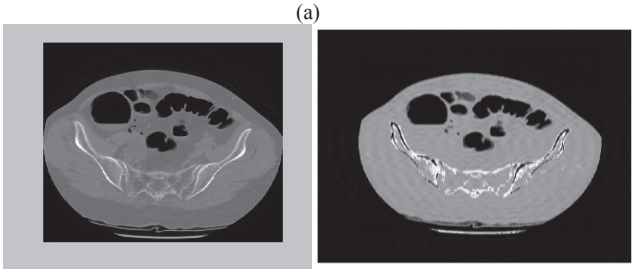
(b) (c)

Figure 1: CT scan slice of the brain (a) Original CT scan slice of the brain (b) SVD transform of the original CT scan slice of the brain image (c)DOST transform of original CT scan slice of the brain image.

The first row of the Fig. 2 shows the original CT abdomen image and the SVD and DOST transforms of the original image are shown in the left and right sides of the second row, respectively. When DOST transform is used, there are losses in the obtained image and the accurate image can not be obtained for CT abdomen image, too.



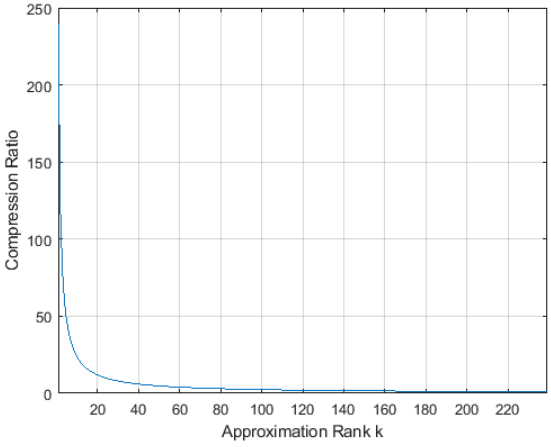
(a)



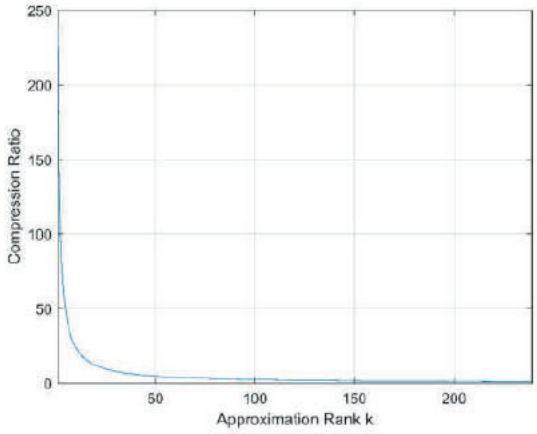
(b) (c)

Figure 2: CT abdomen image. (a) Original CT abdomen image (b) SVD transform of the original CT abdomen image(c) DOST transform of original CT abdomen image.

The reason why SVD transformation performance is better in the analyzed images is that it is possible to create the image with only a few singular values and there have been smaller information loss. However, determination of parameters in DOST technique and the loss of information according to the selected bandwidth and center frequency band values affect the performance of DOST. In addition, the compression ratios on the CT scan slice of the brain and CT abdomen images for SVD transform are depicted in Fig. 3. It can be easily examined from Fig. 3 that for both images, when the rank k is chosen quite less than the number of singular values, the obtained image gets closer to the original image, the compression ratio increases and the storage for the obtained image decreases.



(a)



(b)

Figure 3: Compression ratio for SVD technique (a) Compression ratio for CT scan slice of the brain (b) Compression ratio for CT abdomen image.

In this study, in order to prove the better performance of SVD than DOST technique, MSE and MAE values defined mathematically in (10), are calculated.

$$MSE = \frac{1}{MN}\sum_{i=1}^{M}\sum_{J=1}^{N}(x(i,j)-y(i,j))^2$$
$$MAE = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\left|(x(i,j)-y(i,j))\right| \quad (10)$$

Here, the original image and the reconstructed image are defined as $x(i,j)$ and $y(i,j)$, respectively. In addition, (i,j) is the pixel motion of the $m \times n$ image.

It can be seen from Table 1 that when SVD is used for image compression in medical imaging, better results are obtained and the calculated values for both MSE and MAE get smaller.

Table 1: MAE and MSE valuesfor SVD and DOST techniques

	CT scan slice of the brain image		CT abdomen image	
	SVD	**DOST**	**SVD**	**DOST**
MAE	0.0026	0.0241	2.4187e-05	0.1172
MSE	4.1150e-05	0.0042	9.6422e-08	0.0780

IV. RESULTS AND DISCUSSION

Image compression is an important issue for medical imaging to efficient storage, management of the data size, progressive transmission and image quality. In this study, we compare the SVD and DOST techniques, which are recommended to be used frequently in the existing studies for image compression. In this context, two different CT images are examined and the image quality is calculated by means of MAE and MSE values. As seen in the performance results for both methods, SVD gives better results than DOST. The difference between the original and reconstructed images gets smaller when the SVD technique is applied for image compression. The reason is that by the use of SVD, the image portions are prioritized and the key part of the unit eigenvectors are generated by the most important singular values. Hereby, the first meaningful a few singular values are used and large portion is eliminated. As a result, by using SVD with the optimum k value, the loss of information on the image is decreased and high quality images can be obtained. On the other hand, if the required frequency band and bandwidth values for DOST are not chosen accurately, the image can be distorted.

REFERENCES

[1] Y P Kumar, and K Mishra, " A survey of contemporary medical image compression techniques," Term Project Report, Dept Electronics and Communication Engineering, IIT Guwahati
[2] R Agarwal, C S Salimath, and K Alam, "Multiple image compression in medical imaging technique using wavelets for speedy transmission and optimal storage," Biomedical And Pharmacology Journal, vol 12, pp 183–198, Mar 2019
[3] P P Babu, T J Prasad and K Soundararajan, "Image Compression and Reconstruction using DOST and DCST," Helix, vol 10, pp 77-81, 2020
[4] Y Sain, "Review on Compression of Medical Images using Various Techniques," International Journal of Engineering Research & Technology (IJERT), vol 3 , pp 1290- 1294, August – 2014
[5] M F Ukrit, A Umamageswari, and G R S Suresh, "A survey on lossless compression for medical images," International Journal of Computer Applications, Vol 31, pp 47-50, 2011
[6] S Suma, and V Sridhar, "A review of the effective techniques of compression in medical image processing," International Journal of Computer Applications, vol 97, pp 23-30,July 2014
[7] L Cao, "Singular value decomposition applied to digital image processing," Division of Computing Studies Arizona State University Polytechnic Campus, Mesa, Arizona 85212, pp 1 – 15
[8]] Y Wang, J Orchard, "On the use of the Stockwell transform for image compression, " Proc SPIE 7245 Image Processing: Algorithms and Systems VII, pp 7245-724504, 2009
[9] F Liu, M Hernandez-Cabronero, V Sanchez, M W Marcellin, and A Bilgin, "The current role of image compression standards in medical imaging," Information, vol 8, pp 1-26, 2017
[10] D A Koff, and H Shulman, "An overview of digital compression of medical images: Can we use lossy image compression in radiology? ," Can Assoc Radiol J , Vol 57, pp 211-217, Oct 2006
[11] U Sara, M Akter, and M S Uddin, "Image quality assessment through FSIM, SSIM, MSE and PSNR—A comparative study, " Journal of Computer and Communications, vol 7, pp 8-18, 2019,
[12] D Asamoah, E O Oppong, S O Oppong, and J Danso, "Measuring the performance of image contrast enhancement technique, " International Journal of Computer Applications, vol 181, Oct 2018
[13] B Jahne, "Digital Image Processing" Springer, 2002
[14] R C Gonzalez, R E Woods, and S L Eddins, "Digital image processing using MatLab, " Prentice Hall, 2006
[15] R G Stockwell, L Mansinha, and R P Lowe, "Localization of complex spectrum: the S transform, " IEEE Transactions on Signal Processing, vol 144, pp 998–1001, April 1996
[16] M Eramian, R A Schincariol, R G Stockwell, R P Lowe, and L Mansinha, "Review of applications of 1D and 2D S-transforms, " Proc SPIE3078, pp 558–568, April 1997
[17] R G Stockwell, "A basis for efficient representation of the S-transform, " Digital Signal Processing, vol 17, pp 371–393, January 2007
[18] J M Wardlaw, RADIOLOGY OF STROKE, J Neurol Neurosurg Psychiatry 2001;70(suppl I) i7–i11

[19] M K Kalra, M M Maher, M A Blake, B C Karau, T L Toth, G Avinash, E F Halpern, and S Saini, "Detection and characterization of lesions on low-radiation-dose abdominal CT images post processed with noise reduction filters," Radiology vol 232, pp 791-797, 2004

A survey on Kalman Filtering for Unmanned Aerial Vehicles: Recent Trends, Applications, and Challenges

N.EMER[1] and N.S.ÖZBEK[1]

[1] Department of Electrical and Electronics Engineering, Adana Alparslan Türkeş Science and Technology University, Adana/Turkey, email: nurtenemer2505@gmail.com, nozbek@atu.edu.tr

Abstract— **This study presents recent trends, challenges, applications, and design methodologies of Kalman filter which becomes a cornerstone for the application of unmanned systems. The elaborated filters are range from Kalman to its improved modifications. The algorithms are also widely used in control theory and this aspect of the study makes it a beneficial guide to a large number of readers. Furthermore, the inertial measurement unit, which is the core for the unmanned vehicles, can be designed with different filtering strategies. The filters are compared via a number of features.**

Keywords—Kalman filters, state estimation, inertial measurement unit, fault-detection, unmanned systems.

I. INTRODUCTION

The theoretical and practical improvements of information systems include hardware and software have been essential in the development of unmanned systems [1]. Among them, unmanned aerial vehicles (UAVs) have attracted a great deal of attention in the control-engineering framework. Several advantages have been reported for protecting human life in various missions. To mention a few, UAVs can be useful for many dangerous environments. With this motivation in mind, the potential of UAVs has been increasing day by day in various application areas.

From a safe operating point of view, an unmanned vehicle has to provide all the necessary conditions for reliability, safety, and security. However, the blending of overall imperfections with real environments, measurements, control algorithms, and communication constraints, makes the problem very complicated.

There are various critical steps that must be conducted in the development of UAVs. The fundamental steps can be listed as system identification, sensor fusion, filtering issues, deployment of communication infrastructure, and controller design.

Kalman Filter, which is an estimation algorithm in linear state-space models, has a very important place in linear filter theory [1]. Kalman filter is a recursive estimation filter, which can estimate the current state from the last previously estimated state and new measurement.

The main motivation of this research is the necessity of evaluating Kalman Filter applications for UAVs to express their advantages and drawbacks. Kalman Filter, which is mostly used in many fields along with aviation applications, can produce optimized predictions for the next states according to the previous states of the system. It differs in this from the algorithms of type "batch", because it does not keep any history of measurements or estimates.

In this study, the applications of Kalman filter and its modifications, which have been devoted to several missions of UAV design and control, are investigated.

There are two fundamental problems encountered in practice regarding the mathematical modelling of the systems. The model cannot represent the real system exactly due to the uncertainties. The measurements from the sensors are not perfect. Kalman Filter processes the data obtained by using measurement dynamics, noise in the data, and initial values of the system and performs state estimates of the system with minimum error.

Numerous examples have been reported on the application of Kalman filtering for various tasks in unmanned ground/aerial vehicles navigation [2], motion planning, trajectory estimation and optimization [3], target tracking, guidance and control [4], fault estimation, attitude determination, state estimation [5], and other tasks in the industry [6].

The contributions of this study are given as follows: Various filtering methods used in the design of the inertial measurement unit, which is the core of unmanned aerial vehicle applications, are examined. In addition, general information about the Kalman filter as well as its improvements are given and performance comparisons are presented for each type of Kalman filter. The application examples of Kalman Filter in linear or non-linear systems are detailed with their advantages and drawbacks. From this aspect, the present research provides a practical guide of Kalman filters for unmanned aerial vehicles.

This paper is organized as follows. The theoretical background of Kalman filter is given in Section II, Improvements on Kalman filter are given in Section III. Kalman filter applications for unmanned aerial vehicles are given in Section IV. Finally, the concluding remarks are addressed in the last section.

II. THEORITICAL BACKGROUND OF KALMAN FILTER (KF)

In standard Kalman filter, which is derived from the least-squares method, it is aimed to combine mathematics and the physical world. In control theory, the Kalman filter is known as a linear-quadratic estimator (LQE) that uses a series of measurements, including statistical noise and other errors.

This section investigates the central notion of the Kalman filter. The scheme of data-driven modelling is depicted in Figure 1.

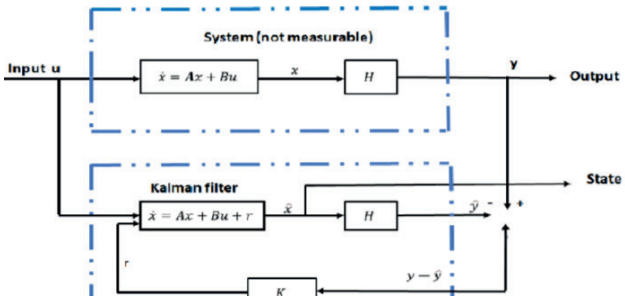


Fig.1. The structure of Kalman Filter

The algorithm takes place in two phases: the first one is the predictive phase, wherein the filter retrieves the last value of the estimated state to estimate the state at the current instant. The second one is the update phase, wherein a new measured value is introduced, allowing to correct the estimation of the first phase, and thus obtain a more precise state.

It is to be noted that the discrete versions of *A*, *B,* and *H* matrices, which are given in Figure 1, are denoted by *F*, *B*, and *H*, respectively. The process model describes the change in state from time *k-1* to time *k* as follows [7]:

$$x_k = F\,x_{k-1} + B\,u_{k-1} + \ w_{k-1} \tag{1}$$

The state transition matrix is defined as *F*, the control input matrix is defined as *B*, previous state vector is defined as x_{k-1}, control vector is defined as $u_{k-1,}$ process noise vector that is assumed to be zero-mean Gaussian with the covariance *Q* is defined as w_{k-1}.

A measurement model should be obtained that describes the relationship between state and measurement in the current time step *k*. This model can be defined as:

$$z_k = Hx_k + \ v_k \tag{2}$$

The measurement vector is defined as z_k, the measurement matrix is defined as *H,* v_k is the measurement noise vector.

Prediction and update are the two structures that compose the algorithm of the Kalman filter [7]. There are two equation in prediction part. The hat operator, "^", is used to mean an estimate of a variable. Superscripts "-" and "+" are used to mean predicted and updated estimates. First one is $\hat{x}_k^-$ which is called as predicted state estimate. Second is P_k^- which is called as predicted error covariance.

$$\hat{x}_k^- = F\,\hat{x}_{k-1}^+ + B\,u_{k-1} \tag{3}$$

$$P_k^- = FP_{k-1}^+F^T + Q \tag{4}$$

There are four equation in update part.

$$\hat{y}_k = \ z_k - H\,\hat{x}_k^- \tag{5}$$

Measurement residual $\hat{y}_k$ is the difference between the true measurement z_k, and the estimated measurement $H\,\hat{x}_k^-$. Calculating Kalman gain K_k is a very important process, because it plays an important role in calculation of update state estimation "$\hat{x}_k^+$".

$$K_k = P_k^-\,H^T\,[R + HP_k^-\,H^T]^{-1} \tag{6}$$

$$\hat{x}_k^+ = \ \hat{x}_k^- + K_k\hat{y}_k \tag{7}$$

$$P_k^+ = (\mathrm{I} - K_kH)\,P_k^- \tag{8}$$

It is important that the updated error covariance "P_k^+" is smaller than the predicted error covariance. This means that the filter is more accurate than the state estimate after it is used in the measurement update phase.

III. Improvements on Kalman Filter (KF)

In real-time applications, many different factors reducing the performance of the Kalman filter have been reported. To mention a few, model errors, as well as model uncertainties unavoidably degrade the performance of Kalman filter. Furthermore, a number of enhancements are required for the Kalman filter to be applied to nonlinear systems. Thus, a number of improvements and modifications have been investigated on the application of Kalman Filter such as Extended Kalman Filter (EKF), Unscented Kalman Filter (UKS), Robuts Kalman Filter (RKF), Adaptive Kalman filter (AKF).

A. Extended Kalman Filter (EKF)

The extended Kalman filter is one such example for these enhancements. EKF is a filter created by replacing the traditional Kalman filter, especially for the application of system or measurement models in non-linear situations. The EKF is a filter that considers continuous-time dynamic systems and discrete-time measurements for nonlinear system estimation through linearization. EKF uses a linearization process ignoring high-order terms in the Taylor series and requiring Jacobian matrix calculation [2, 4].

$$\dot{x} = f(x,u,t) + D(t)\,w \tag{9}$$

$$z_k = y\,[\,x(t_k), k\,] + \ v_k \tag{10}$$

where a system model is defined as $f(x,u,t)$ and a measurement model is defined as $y(x,k)$. Further, w and v_k are the process and measurement noises, respectively.

$$w \approx (0,Q)\ ,\ v_k \approx (0,R)$$

The first step in prediction is to predict the next state of the system input *u*.

$$\dot{\hat{x}} = f(\,\hat{x}, u, t) \tag{11}$$

$$\dot{P} = F(\,\hat{x}\,, t\,)P + PF^T\,(\,\hat{x}\,, t\,) + DQD^T \tag{12}$$

$f(\,\hat{x}, u, t)$ and P can be obtained from prior measurement update and the measurement estimate is updated as follows [4]:

$$K_k = P^-\,(\,t_k)\,Y^T(\,\hat{x}_k^-)[Y\,(\,\hat{x}_k^-)\,P^-(\,t_k)\,Y^T\,(\,\hat{x}_k^-)\ + R\,]^{-1} \tag{13}$$

$$P(t_k) = [\,I - K_k\,Y(\,\hat{x}_k^-)\mathrm{P}^-(t_k) \tag{14}$$

$$\hat{x}_k = \ \hat{x}_k^- + K_k[\,z_k - Y(\,\hat{x}_k^-\,, k)] \tag{15}$$

$F(x,t)$ and $Y(x)$ are Jacobians derived from $f(\,\hat{x}, u, t)$ and $y(x, k)$ [4].

$$F(\,x, t\,) = \frac{\partial \mathrm{f}(\,x,u,t)}{\partial x}\ ,\ Y\,(\mathrm{x}) = \frac{\partial \mathrm{y}(\,x,k)}{\partial x} \tag{16}$$

B. Unscented Kalman Filter (UKF)

In Extended Kalman Filter, linearization is performed using Jacobian matrix and calculations of Jacobian matrix increase the cost. Moreover, when estimating in high-order nonlinear systems in Extended Kalman Filter, it is difficult to obtain correct results because due to the linearization. In order to reduce these disadvantages, EKF-based Unscented Kalman Filter (UKF) was developed [1] .

UKF captures mean and covariance estimates with a deterministic sampling approach [8]. Thus, the UKF is easy to implement. Furthermore, UKF provides more robust forecast performance.

C. Robust Kalman Filters (RKF)

The Robust Kalman filter can be applied in uncertain discrete-time systems [9]. It provides an upper-bound to the variance of the filtering error to resolve parameter uncertainties in the state and output matrices [10]. A robust Kalman filter (RKF) processing data under noisy measurements has been applied to the FPID control strategy. Thus, the robustness and anti-interference feature of the UAV system is increased and the effect of outliers on the updated parameters is minimized [22]. A further application of RKF with a Gaussian uniform mixture model is presented in [13].

D. Adaptive Kalman Filter(AKF)

In the classical Kalman filter, an accurate system model and stochastic information are required for predictions to be correct. However, it is very difficult to obtain exact system information in real applications. Adaptive Kalman Filter (AKF) is proposed to reduce the impact of this deficiency. AKF overcomes errors in the prediction of nonlinear state-space models that may result from incomplete system information [14]. An attitude estimation by using an AKF is presented in [15].

IV. Kalman Filter Applications for Unmanned Aerial Vehicles

This section presents the applications of the Kalman Filters for UAV applications. The elaborated filters are applied for several tasks such as system identification, sensor fusion, state estimation, control, fault detection, as well as fault-tolerant control. The application of Kalman filters UAVs is tabulated in Table I.

Table 1: Kalman Filter Applications for UAVs

	KF	EKF	UKF		AKF
Identification	[16], [17]	[18],[19], [20], [21]	[22], [23]		[24]
	[25]	[26]	[27]		
Control Design	[28], [29]	[4], [5], [21],[30]	[31], [32]		[33]
Fault-Tolerant Control					

A. Kalman Filters for Identification of UAV

System identification, which can be used to characterize different components in unmanned systems or to construct the mathematical model of the system with measured input-output data [40], is an important step in designing the controllers of UAVs. In addition, a discrete model for predictive control of UAVs can be obtained by system identification. The obtained models can be used as a prediction system in controllers.

The reports on the Kalman filtering devoted to system identification can be summarized as follows: Kim Y. et al. proposed two filters to make parameter estimates. Kalman filter (KF) for linear dynamics and Unscented Kalman Filter UKF for nonlinear dynamics are proposed[17].

It is to be noted that UAVs have a multi-input-multiple output structure. Thus, data from sensors on UAVs are mostly non-linear and time-varying values. This makes it tedious to estimate the parameters and determine the system model. For nonlinear systems, the classical Kalman filters are evaluated to be inadequate. An EKF has been proposed to define system dynamics in [20]. It is to be noted that EKF, which uses an iterative nonlinear filtering approach, can be applied separately for linear and non-linear models in the identification of UAVs.

A further example of EKF has been proposed for both state and parameter estimation of a UAV [21]. EKF is proposed to deal with nonlinear effects in [18], wherein, determination of the stability and control parameters of the nonlinear unmanned aerial vehicle model is investigated. While making these definitions, the dynamic effects of some external factors occur. Furthermore, Munguía R. et al. propose an improved EKF that estimates the model parameters of the UAV using measurements directly obtained from sensors on multi-rotor UAVs. With these observations, EKF, which updates with the measurements taken from the sensors, can be able to predict all model parameters of multi-rotor aerial vehicles [19].

Information obtained from many sensors can be used during the design of controllers. However, a number of problems are encountered in cases such as noisy measurements as well as the nonlinearity of the data received from the sensors. Some filters are recommended to tolerate these problems. In this article, UKF and EKF performances have been examined and compared for parameter estimation [22]. UAVs are highly complex systems with multi inputs. Thus to construct an exact mathematical model is a challenging issue. However, a UKF is highlighted to estimate the model of the UAV through the flight data [19].

Chiella et al. define the mathematical equivalent model to protect the lithium-ion batteries in UAVs. Further, the parameters used to predict the next charge rate of the battery are estimated. Toward this goal, an adaptive Kalman filtering (AKF) is addressed [24].

B. Kalman Filters for Sensor Fusion of UAV

Sensor fusion and the filtering of measurements are one of the most important points for unmanned vehicles. The investigation of sensor fusion techniques has been very attractive. For these purposes, Kalman filter is a powerful tool for sensor fusion applications with its recursive nature. With this motivation in mind, this section presents filtering issues dedicated to sensor fusion of UAVs.

The performance of the estimation strategies is elaborated with a number of examples. For instance, Arreola L. et. al. propose a position estimation strategy which is based on low-cost devices and optical flow algorithm. Towards this goal, an extended Kalman filter algorithm is employed [25]. A location estimation system using low-cost devices for UAVs has been proposed in [25], wherein the information from GPS and inertial navigation sensors (INS) are used to estimate the position of the UAV. GPS and INS information is fused in the Kalman Filter. In addition, EKF is presented to estimate the position using data obtained from the optical flow algorithm to improve the UAV position estimation [25]. A low-cos PX4Flow optical sensor, which is an open-source hardware and software platform comprising CMOS vision sensor, a gyroscope, and a sonar range finder, is addressed in [41].

Furthermore, a sensor fusion method, which is based on EKF algorithm, is addressed for estimating the velocity, position, and attitude of a UAV with a low-cost inertial measurement unit [26]. A number of experiments are conducted with various sampling intervals and loss data conditions to investigate the proposed algorithm.

An innovative fusion filter-based method is proposed by Yue L. et al to solve the problem of tracking quadrotors in the presence of interference and measurement noise. An example block diagram of a control structure of the quadrotor is shown in Figure 2. Traditional PID cannot properly adjust its gain with environmental distortion. FPID has been proposed to reduce this disadvantage. In order to complement the shortcomings of FPID, the EKF and RKF are proposed. When the Kalman filter (KF) is applied to the system, the presence of outliers can seriously damage the compensation effect. If the necessary suppression measures are not taken, this abnormal value will greatly damage the monitoring performance of the system [11].

Another example of UKF based Kalman filter application is reported by Chiella et. al. [27], wherein a robust adaptive fusion algorithm is proposed for collecting data in an actual forest environment.

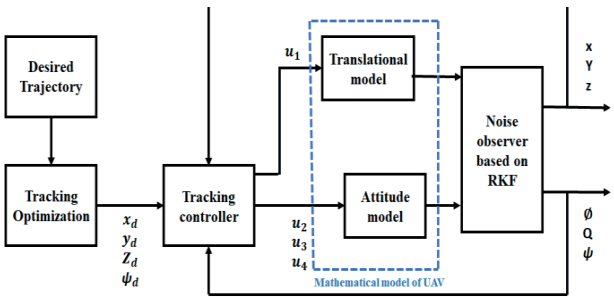


Fig.2. The control system of quadrotor [37]

C. Kalman Filters for State Estimation and Control

The main purpose of using the Kalman filter is to be able to make state-estimation, rate estimation. This filter has increased the reliability of complex systems by reducing noise. Kwon H. et. al. address an EKF based Sliding mode control for attitude stabilization. The proposed algorithm estimates the attitude state by an EKF, then stabilizes the system with a variable structure control [4]. Song Q. et. al. address a control scheme that combines backstepping and PD controller for the position and attitude control, respectively. Further, an unscented Kalman filter for both fast and slow modes estimates a wind-gust disturbance, which is to be considered as an external disturbance [31].

In a further study, an attitude estimation of a quadrotor type aerial vehicle is addressed through an observer that is based on Extended Kalman Filter [5]. The data losses are model as a random process.

Extended Kalman filter is employed separately for linear and non-linear models in the identification of unmanned aerial vehicles. The extended Kalman filter has been proposed for both UAV's state and parameter estimation [21]. The UKF has been recommended where EKF is inadequate to deal with some of the difficulties encountered in making the forecast of unmanned aerial vehicles [32]. Wang D. et al. propose a system to track the target and determine the states of the UAV in non-linear situations. To do this, a new AUKF system has been proposed by combining the UKF and AKF [33].

Jimenez et. al. analysis the total energy control systems in control of a UAV [42]. Furthermore, an EKF approach is addressed to predict angular rates as well as linear acceleration during the test flight. Both techniques are experimentally validated on a fixed-wing UAV. In another study, Zhang K. et al. investigate the flight control and tracking problem of a quadrotor with model uncertainties and external disturbances [30]. A linear-quadratic regulation (LQR) tracking algorithm is addressed, however, several difficulties such as high nonlinearities, coupled dynamics, parameter variations, and uncertainties in the system dynamics, are encountered in LQR design. To overcome these drawbacks, an Extended Kalman Filter (EKF) based LQR is proposed for online estimation of position, speed, and engine dynamics using measured outputs. It is to be noted that more effective results can be obtained for disturbance rejection by the use of EKF [30].

The block diagram of the controller, which is designed based on LQR and EKF techniques, is shown in Figure 3.

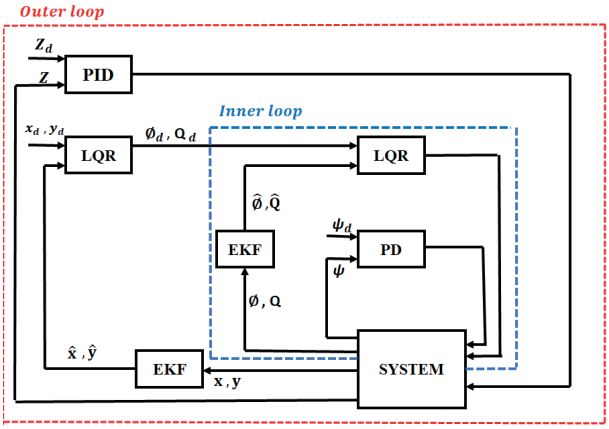


Fig. 3. The inner-outer control structure[30]

Hetenyi D. et. al. propose a method to check the position of the UAV [28]. However, the presence of noise brings a number of difficulties and makes the design of these control systems difficult. Thus, it is necessary to filter the signals from the sensors to create a stable and accurate measurement. With this motivation in mind, the Kalman filter is used to combine the ultrasonic range finders and accelerometer signals. As a result, it has been shown that the error generated by Kalman filter is relatively small and valid for correct altitude application [28].

D. Kalman Filters for Fault-Detection and Fault-Tolerant Control of UAV

Fault detection, diagnostics, and fault-tolerant control are important in developing control algorithms, which can deal with failures occurring in critical tasks. A number of research studies can be summarized as follows: Egidio D'Amato, et. al. propose a fault-tolerant system for attitude estimation of UAV using low-cost magnetometers, accelerometers, and gyroscopes. The Unscented Kalman Filter (UKF) has been used to propose an approach for detection, isolation, and reconstruction. It reduces the effect by identifying the error by comparing the predictions made by the UKF with the measured variables. The disadvantage of the filter is defined as the small delay in isolating the faults in the use of the UKF and the increase in the calculation cost. However, in the presence of a double fault, UKF guarantees a fail-safe behaviour [37].

It is important to determine the state of unmanned aerial vehicles correctly. However, some errors, noises, and distortions may be encountered in the state determinations. In these cases, it is necessary to detect the faults correctly, to eliminate their effects on the system, and to prevent the deterioration of the system characteristics. RKF is suggested for realizing these situations in [14].

Ngo T. et al. address a method to control the vehicles in an unknown outdoor environment. To do this, the complementary filter and Kalman filter are designed to predict the exact state vector showing the movement of the body. Experimental results are presented to validate the design of these filters. The complementary filter takes gyro values as major factors and accelerometer values as minor factors. Therefore, the oscillation reflected in the angles of the accelerometer cannot affect the system much. It follows from this that the estimated angle of the complementary filter is more stable than the Kalman filter [29]. When buildings or trees block the UAV during flight, the GPS receiver cannot receive all satellite signals and generates an incorrect positioning signal that can be expressed as outliers present in the measurement noise. Therefore, it is very important to estimate the UAV's position.

It is known that faults, which have nonlinear nature, may cause fatal damages in unmanned systems. An AKF is proposed by Weimer F. et al. [36] to detect faults as well as to make correct definitions when different faults are encountered. In a further study, an actuator based sensor fault detection, as well as diagnosis algorithm, are presented in [34]. An actuator fault detection and diagnosis algorithm are proposed for a quadrotor type aerial robot in the presence of external disturbances [38]. The proposed strategy is based on an adaptive augmented state Kalman filter. A neural network-based system has been proposed to detect malfunctions in unmanned aerial vehicles. This system is strengthened with EKF. It is also used to update the parameters in the EKF Neural Network [35]. Liang Y. et al. propose EKF to detect the error in the UAV's navigation system and to obtain the correct values in the presence of any error [36].

V. CONCLUSION

The practical improvements of UAVs have been increasing with the development of control system technologies, signal processing methods, and communication capabilities. However, the design, modelling, and control of UAVs bring a number of challenges due to the highly complex nonlinear nature.

The main goal of the current research is to explore Kalman filtering techniques devoted to system identification, sensor fusion, state estimation, fault detection, fault-tolerant control of UAVs. Furthermore, it is aimed to guide the control researchers on the filtering techniques during the modelling and control of critical systems. With this motivation in mind, recent studies on Kalman filtering for UAV applications are addressed.

ACKNOWLEDGMENT

The authors are gratefully acknowledged to anonymous reviewers. This study is financially supported by the Scientific Research Projects Unit of Adana Alparslan Türkeş Science and Technology University under the 20303003 project number.

REFERENCES

[1] M Zhang, K Li, B Hu, and C Meng, "Comparison of kalman filters for inertial integrated navigation," *Sensors (Switzerland)*, vol 19, no 6, 2019

[2] J Farrell and M Barth, *The Global Positioning System and Inertial Navigation* New York, NY, USA,: McGraw-Hill, 1998

[3] B Kada, K Munawar, M S Shaikh, M A Hussaini, and U M Al-Saggaf, "UAV Attitude Estimation Using Nonlinear Filtering and Low-Cost Mems Sensors," *IFAC-PapersOnLine*, vol 49, no 21, pp 521–528, 2016

[4] H Kwon, K Lee, and K You, "EKF based sliding mode control for a quadrotor attitude stabilization," *ICIIBMS 2017 - 2nd Int. Conf. Intell. Informatics Biomed. Sci.*, vol 2018-Janua, pp 101–104, 2018

[5] S Lesecq, S Gentil, and N Daraoui, "Quadrotor attitude estimation with data losses," *2009 Eur. Control Conf. ECC 2009*, pp 3851–3856, 2014

[6] H Ma, L Yan, Y Xia, and M Fu, *Kalman Filtering and Information Fusion* Singapore: Springer Singapore, 2020

[7] Y Kim and H Bang, "Introduction to Kalman Filter and Its Applications," *Introd. Implementations Kalman Filter*, pp 1–16, 2019

[8] A Akca and M Ö Efe, "Multiple Model Kalman and Particle Filters and Applications: A Survey," *IFAC-PapersOnLine*, vol 52, no 3, pp 73–78, 2019

[9] M Zorzi, "Robust Kalman Filtering Under Model Perturbations," *IEEE Trans. Automat. Contr.*, vol 62, no 6, pp 2902–2907, Jun 2017

[10] Z Wang, "Robust filter design with time-varying parameter uncertainty and error variance constraints," *Int. J. Control*, vol 72, no 1, pp 30–38, Jan 1999

[11] Y Li, Z Wang, S Deng, and Q Ouyang, "Tracking control of a quadrotor UAV based on fusion filter," *Proc. - 2019 Chinese Autom. Congr. CAC 2019*, pp 2641–2645, 2019

[12] N S Ozbek, "An Evaluation of Model-Free Control Strategies for Quadrotor Type Unmanned Aerial Vehicles," *Proc. - 2019 3rd Int. Conf. Appl. Autom. Ind. Diagnostics, ICAAID 2019*, no September, pp 25–27, 2019

[13] M Brunot, "A Gaussian Uniform Mixture Model for Robust Kalman Filtering," *IEEE Trans. Aerosp. Electron. Syst.*, vol 56, no 4, pp 2656–2665, Aug 2020

[14] C Hajiyev and H E Soken, "Robust Adaptive Kalman Filter for estimation of UAV dynamics in the presence of sensor/actuator faults," *Aerosp. Sci. Technol.*, vol 28, no 1, pp 376–383, Jul 2013

[15] Y Luo, G Ye, Y Wu, J Guo, J Liang, and Y Yang, "An Adaptive Kalman Filter for UAV Attitude Estimation," in *2019 IEEE 2nd International Conference on Electronics Technology (ICET)*, 2019, pp 258–262

[16] Y Xu, Y S Shmaliy, X Chen, Y Li, and W Ma, "Robust inertial navigation system/ultra wide band integrated indoor quadrotor localization employing adaptive interacting multiple model-unbiased finite impulse response/Kalman filter estimator," *Aerosp. Sci. Technol.*, vol 98, p 105683, Mar 2020

[17] G Seo, Y Kim, and S Saderla, "Kalman-filter based online system identification of fixed-wing aircraft in upset condition," *Aerosp. Sci. Technol.*, vol 89, pp 307–317, Jun 2019

[18] C Grillo and F Montano, "An Extended Kalman Filter-based technique for on-line identification of unmanned aerial system parameters," *J. Aerosp. Technol. Manag.*, vol 7, no 3, pp 323–333, 2015

[19] R Munguía, S Urzua, and A Grau, "EKF-based parameter identification of multi-rotor unmanned aerial vehiclesmodels," *Sensors (Switzerland)*, vol 19, no 19, pp 1–17, 2019

[20] A G Kallapur, S S Ali, and S G Anavatti, "Application of extended Kalman filter towards UAV identification," *Stud. Comput.*

Intell., vol 76, pp 199–207, 2007
[21] A G Kallapur and S G Anavatti, "UAV linear and nonlinear estimation using extended kalman filter," *CIMCA 2006 Int. Conf. Comput. Intell. Model. Control Autom. Jointly with IAWTIC 2006 Int. Conf. Intell. Agents Web Technol. ...*, 2006
[22] G Chowdhary and R Jategaonkar, "Aerodynamic parameter estimation from flight data applying extended and unscented Kalman filter," *Aerosp. Sci. Technol.*, vol 14, no 2, pp 106–117, Mar 2010
[23] B Wang, D Liu, W Wang, and X Peng, "A hybrid approach for UAV flight data estimation and prediction based on flight mode recognition," *Microelectron. Reliab.*, vol 84, pp 253–262, May 2018
[24] S Wang *et al.*, "A novel safety assurance method based on the compound equivalent modeling and iterate reduce particle-adaptive Kalman filtering for the unmanned aerial vehicle lithium ion batteries," *Energy Sci. Eng.*, vol 8, no 5, pp 1484–1500, May 2020
[25] L Arreola, A Montes De Oca, A Flores, J Sanchez, and G Flores, "Improvement in the UAV position estimation with low-cost GPS, INS and vision-based system: Application to a quadrotor UAV," *2018 Int. Conf. Unmanned Aircr. Syst. ICUAS 2018*, pp 1248–1254, 2018
[26] S P H Driessen, N H J Janssen, L Wang, J L Palmer, and H Nijmeijer, "Experimentally Validated Extended Kalman Filter for UAV State Estimation Using Low-Cost Sensors," *IFAC-PapersOnLine*, vol 51, no 15, pp 43–48, 2018
[27] A C B Chiella, B O S Teixeira, and G A S Pereira, "State estimation for aerial vehicles in forest environments," *2019 Int. Conf. Unmanned Aircr. Syst. ICUAS 2019*, pp 890–898, 2019
[28] D Hetényi, M Gótzy, and L Blázovics, "Sensor fusion with enhanced Kalman Filter for altitude control of quadrotors," *SACI 2016 - 11th IEEE Int. Symp. Appl. Comput. Intell. Informatics, Proc.*, pp 413–418, 2016
[29] T Ngo, P Nguyen, S Huynh, S Le, and T Nguyen, "Filter Design for Low-Cost Sensor in Quadcopter," *2017 Int. Conf. Syst. Sci. Eng.*, no 3, pp 488–493, 2017
[30] K Zhang, J Chen, Y Chang, and Y Shi, "EKF-based LQR tracking control of a quadrotor helicopter subject to uncertainties," *IECON Proc. (Industrial Electron. Conf.*, pp 5426–5431, 2016
[31] Q Song and S Ji, "UKF Based Two Time-Scale Control for a Quadrotor with Wind-Gust Disturbances," *2018 3rd Int. Conf. Robot. Autom. Eng. ICRAE 2018*, pp 191–195, 2018
[32] H Bonyan Khamseh, S Ghorbani, and F Janabi-Sharifi, "Unscented Kalman filter state estimation for manipulating unmanned aerial vehicles," *Aerosp. Sci. Technol.*, vol 92, pp 446–463, Sep 2019
[33] D Wang, H Lv, and J Wu, "In-flight initial alignment for small UAV MEMS-based navigation via adaptive unscented Kalman filtering approach," *Aerosp. Sci. Technol.*, vol 61, pp 73–84, Feb 2017
[34] M Moghadam and F Caliskan, "Actuator and sensor fault detection and diagnosis of quadrotor based on Two-Stage Kalman Filter," *2015 Aust. Control Conf. AUCC 2015*, no June, pp 182–187, 2015
[35] A Abbaspour, P Aboutalebi, K K Yen, and A Sargolzaei, "Neural adaptive observer-based sensor and actuator fault detection in nonlinear systems: Application in UAV," *ISA Trans.*, vol 67, pp 317–329, Mar 2017
[36] Y Liang and Y Jia, "A nonlinear quaternion-based fault-tolerant SINS/GNSS integrated navigation method for autonomous UAVs," *Aerosp. Sci. Technol.*, vol 40, pp 191–199, Jan 2015
[37] E D'Amato, M Mattei, A Mele, I Notaro, and V Scordamaglia, "Fault tolerant low cost IMUS for UAVs," *2017 IEEE Int. Work. Meas. Networking, M N 2017 - Proc.*, pp 0–5, 2017
[38] Y Zhong, Y Zhang, W Zhang, J Zuo, and H Zhan, "Robust Actuator Fault Detection and Diagnosis for a Quadrotor UAV with External Disturbances," *IEEE Access*, vol 6, pp 48169–48180, 2018
[39] F Weimer, T Rothermel, and W Fichter, "Adaptive Actuator Fault Detection and Identification for UAV Applications," *IFAC Proc. Vol.*, vol 45, no 1, pp 67–72, 2012
[40] N S Ozbek and I Eker, "A novel interactive system identification and control toolbox dedicated to real-time identification and model reference adaptive control experiments," in *2015 9th International Conference on Electrical and Electronics Engineering (ELECO)*, 2015, pp 859–863
[41] D Honegger, L Meier, P Tanskanen, and M Pollefeys, "An open source and open hardware embedded metric optical flow CMOS camera for indoor and outdoor applications," *Proc. - IEEE Int. Conf. Robot. Autom.*, pp 1736–1741, 2013
[42] P Jimenez, P Lichota, D Agudelo, and K Rogowski, "Experimental Validation of Total Energy Control System for UAVs," *Energies*, vol 13, no 1, p 14, Dec 2019

Analysis of Foundations Settled On Different Soil Classes by Considering Different Soil Models

B. KAYABAŞ[1] and H. F. PULAT[1]

[1]İzmir Katip Çelebi University, İzmir/Turkey, kayabasbuket@hotmail.com
[1] İzmir Katip Çelebi University, İzmir/Turkey, hfirat.pulat@ikcu.edu.tr

Abstract **- During the design of the foundations, many parameters such as soil class, soil properties, drainage conditions, type of structure should be taken into consideration. In case these parameters are not examined sufficiently, uneconomical, overdesigned foundations can be constructed. For this purpose, different soil models such as Mohr-Coulomb (MC), Hardening Soil (HS), Soft Soil (SS) have been developed in the evaluation of geotechnical engineering. On the other hand, various soil mechanic theories were developed to define the bearing capacity of the soils (Terzaghi, Meyerhof, Hansen, Vesic, etc.).**

In this study, five projects constructed on different soil classes were analyzed with Plaxis. Moreover, the bearing capacities of the different soil classes were calculated with extensively used bearing capacity theories. By comparing the analysis outputs with the bearing capacity theories, more consistent and reasonable results were determined.

Keywords **- Bearing capacity theories, finite element method, soil models.**

I. INTRODUCTION

Foundation design in geotechnical engineering should be made by considering many parameters. One of these parameters and the most important is the soil models used in the analysis (Linear elastic, Mohr-Coulomb, Soft soil, Soft soil creep, hardening soil, jointed-rock model etc.). On the other hand, the bearing capacity of the soil plays a critical role in the design process. Bearing capacity theories that differ from each other in many ways have been developed (Terzaghi, Meyerhof, Hansen, Vesic, Prandtl, Bowles, Parry, Pressiometer etc.). These parameters are important for getting more accurate results in analysis and making more economical designs.

Ercan (2001), based on Krinitzsk's approach to static bearing capacity of the ground for building foundations, developed the bearing capacity and soil safety stress relationship accordingly. It has been determined that the results obtained from the approach are consistent with the laboratory results [5].

Castaldo et al. (2014) investigated the seismic properties of reinforced concrete structures around deep excavations before and after excavation. For analysis, a finite element based program that takes into account the soil-structure interaction was used. The behavior of the soil against seismic effects was determined using HS and MC models. As a result of the analysis, significant differences were determined between the seismic behavior of the reinforced concrete structure before and after the excavation [3].

Hsiung et al., (2016) examined the excavation support systems built in an area of thick sand layer in Kaohsiung, Taiwan, with various computer models. The results of the analysis performed with the HS, HS (SS) and MC models in the PLAXIS program were compared with the field results [7].

In this study, five different projects have been analyzed using different soil models. Analyzes have been made by considering projects with different engineering characteristics and soil conditions. Bearing capacities of soils have been calculated using different theories. Analysis and calculation results were compared and plausible result pairs were determined.

II. METHOD

A. Bearing Capacity Theories

Terzaghi's bearing capacity theory; Terzaghi (1943) was concerned with the bearing strength of rough, rigid, shallow and varietal foundations resting on homogeneous ground [11]. The sliding surface of this method is given in Figure 1. Here, only bearing capacity factors that vary depending on the internal friction angle of the ground are used.

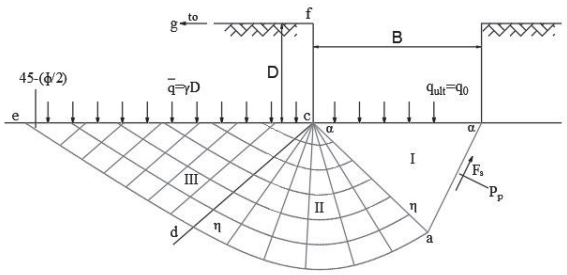


Figure 1. The sliding surface of a continuous, rough, rigid foundation [8]

The equation Terzaghi uses;

$$q_n = c \cdot N_c \cdot s_c + \sigma_z \cdot N_q + 0.5 \cdot \gamma \cdot B \cdot N_\gamma \cdot s_\gamma \quad (1)$$

Meyerhof's bearing capacity theory; unlike the Terzaghi theory, Meyerhof (1951) includes the part of the sliding surface that remains above the foundation base in the calculation [9]. In

addition, the angle made by the wedge with the horizontal axis is equal to 45+ϕ/2. The foundation base must frictional as in the Terzaghi method. Meyerhof developed Terzaghi's equations with shape, depth, slope factors and obtained the following equation.

$$q_n = c \cdot N_c \cdot s_c \cdot d_c + \sigma_z \cdot N_q \cdot s_q \cdot d_q + 0.5 \cdot \gamma \cdot B \cdot N_\gamma \cdot s_\gamma \cdot d_\gamma \quad (2)$$

Hansen's bearing capacity theory; Hansen (1961) developed the formula by adding load slope angle factors, soil slope factors and fundamental deviation factors to Meyerhof's formulas [6]. Hansen bearing strength formula is as follows.

$$q_n = c \cdot N_c \cdot s_c \cdot d_c \cdot i_c \cdot g_c \cdot b_c + \sigma_z \cdot N_q \cdot s_q \cdot d_q \cdot i_q \cdot g_q \cdot b_q + 0.5 \cdot \gamma \cdot B \cdot N_\gamma \cdot s_\gamma \cdot d_\gamma \cdot i_\gamma \cdot g_\gamma \cdot b_\gamma \quad (3)$$

Vesic's bearing capacity theory; although Vesic (1973) bearing capacity calculation method is similar to the bearing capacity method developed by Hansen, some parameters vary [12]. Vesic did not take into account the eccentricity effect in his calculations and accepted the basic dimensions as is.

Prandtl's plastic equilibrium theory;

The Prandtl plastic equilibrium theory formula is as follows.

$$q_u = \left[c \cdot \cot\phi + \frac{\gamma B}{2} \tan\left(45 + \frac{\phi}{2}\right) \right] \left[\tan^2\left(45 + \frac{\phi}{2}\right) e^{\pi \tan\phi} - 1 \right] \quad (4)$$

Bowles's bearing capacity calculation; this method assumes that the basic bearing capacity calculated according to the SPT-N numbers are obtained by staying on the very safe side. Equations used for this method;

$$q_e = \frac{N}{F_1} \cdot K_d \qquad B < F_4 \quad (5)$$

$$q_e = \frac{N}{F_2} \cdot \left(\frac{B + F_3}{B} \right)^2 K_d \qquad B > F_4 \quad (6)$$

$$K_d = 1 + 0.33 * \frac{D}{B} \leq 1.33 \quad (7)$$

Parry's bearing capacity calculation; Parry (1977) referred to the safe bearing capacity calculation for noncohesive granular floors [10]. The equation used for this method;

$$q_e = 30 * N_{55} \quad (8)$$

Bearing capacity calculation with pressiometer Experiment; with this experiment, the ultimate bearing capacity of the soil is calculated as follows.

$$q_n = \sigma_z + k_{bc} * \left(p_{maks} - \sigma_h \right) \quad (9)$$

B. Soil Models

Linear elastic model (LE); is the most common material model. The linear elastic model is insufficient in solutions because soils show nonlinear behavior. In this model, since the stress state is not restricted, infinite stiffness occurs in the material and the model behaves out of reality.

Mohr–Coulomb model (MC); examines the stress-strain curve of the soil in two regions. In the first region of this graph, the soil is assumed to be elastic. In its second zone, the floor is considered to be perfectly plastic. In this model, the stresses are considered to be equal to the values of the soil at the moment of failure criteria. Therefore, this model is used in the preliminary analysis of systems. In this model, it is assumed that the stiffness of the soil layers does not differ with depth. Usually; It is used in the safety analysis of retaining walls, earthfill dams, slopes, and is used in the bearing capacity of shallow foundation projects.

Soft soil model (SSM); The soft soil model was developed by Brinkgreve, R.B.J. (2002) [2]. This model is based on the Mohr-Coulomb criterion on the yield surface. This model is used on soft soils with high compressibility and normally consolidated soils under primary compression conditions.

Soft soil creep model (SSCM); It is used on soft soils with different properties, which are clayey silts, peat soils and normally consolidated clay. Considering the creep behavior in this model, it is assumed that the elastic deformation is instantaneous, the plastic deformation is only viscous and will develop over time.

Hardening soil model (HS); this model developed by Duncan and Chang (1970) is suitable for different types of soft and hard soils [4]. Although this model is considered suitable for all soil types, it is not used on such soils because it cannot take into account the viscous effects such as shrinkage-stress relaxation in soft soils and the time-dependent compression behavior in normally consolidated clays. This model will give near accurate results for gravels, sands, silts and clays.

Jointed – Rock Model (JR); It is an anisotropic elastoplastic model in which plastic slip occurs only in a limited number of slip directions. This model is often used when modeling bedded rocks and compound rocks.

C. Finite Element Method

Plaxis program is a finite element program used in geotechnical engineering projects and provides a rapid generation of mixed finite element models. The finite element method is an approximate method and realistic results can be obtained with sufficient elements. It is necessary to create a model for each project first. There are three main types of components found in a model geometry. These; dots, lines, and cells.

In addition, the Plaxis package program consists of four main sections.

- Input
- Calculation
- Output
- Curves

In the modeling phase, first of all, the properties of the soil layers are determined. Then the load on the foundation and foundation is given. After these are completed, mesh is applied to the soil. After the groundwater level is determined, the solution window is passed. Steps should be created for the solution in the calculation window. After the stages are created, the solution is started. If there is no problem in the solution, the graphic reading process is started. After the calculation, axial force, shear force, moment value and total displacement results are obtained. With the values read from these graphs obtained, the stresses occurring at the foundation can be calculated.

III. GENERAL INFORMATIONS ABOUT PROJECT

I. Project Data

In this project, Mohr-Coulomb Model and Hardening Soil Model were used considering the soil profiles. The idealized soil profile and information about the building are given below.

Table 1. Building information

Province	İzmir
District	Menemen
Base area	780 m^2
B	20 m
L	39 m
D_f	3.6 m
Floor	9
Total height	22.00 m

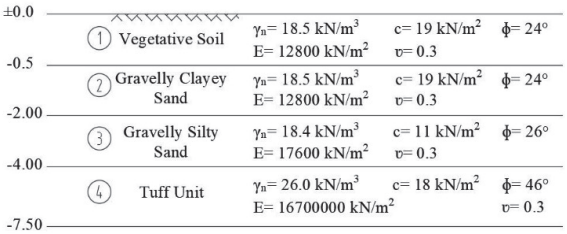


Figure 2. Idealized soil profile

II.Project Data

In this project, Mohr-Coulomb Model and Hardening Soil Model were used considering the soil profiles. The idealized soil profile and information about the building are given below.

Table 2. Building information

Province	İzmir
District	Buca
Base area	1200 m^2
B	20 m
L	60 m
D_f	17.6 m
Floor	14
Total height	29.50 m

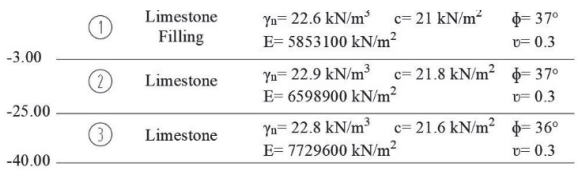


Figure 3. Idealized soil profile

III.Project Data

In this project, Mohr-Coulomb Model and Hardening Soil Model were used considering the soil profiles. The idealized soil profile and information about the building are given below.

Table 3. Building information

Province	İzmir
District	Torbalı
Base area	1560 m^2
B	26 m
L	60 m
D_f	5.6 m
Floor	9
Total height	24.00 m

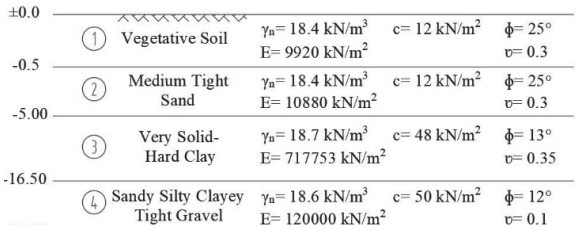


Figure 4. Idealized soil profile

IV.Project Data

In this project, Mohr-Coulomb Model, Soft Soil Model and Hardening Soil Model were used considering the soil profiles. The idealized soil profile and information about the building are given below.

Table 4. Building information

Province	İzmir
District	Bornova
Base area	800 m^2
B	20 m
L	40 m
D_f	3.52 m
Floor	11
Total height	32.00 m

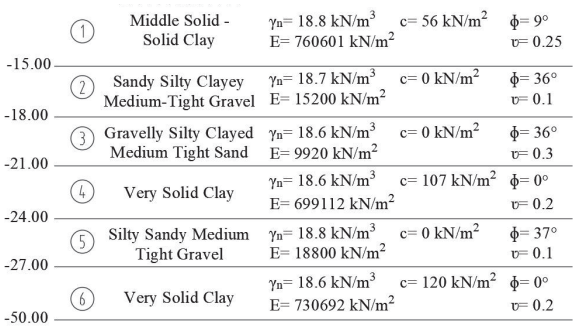


Figure 5. Idealized soil profile

V.Project Data

In this project, Mohr-Coulomb Model, Soft Soil Model and Hardening Soil Model were used considering the soil profiles. The idealized soil profile and information about the building are given below.

Table 5. Building information

Province	İzmir
District	Konak
Base area	400 m^2
B	20 m
L	20 m
D_f	7.17 m
Floor	11
Total height	30.70 m

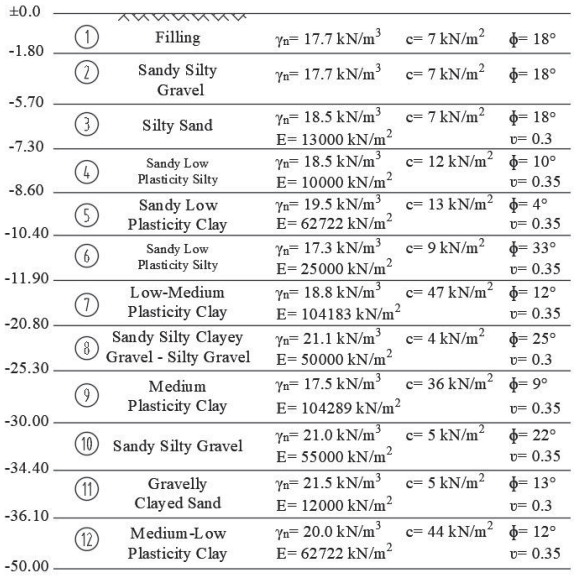


Figure 6. Idealized soil profile

IV. ANALYSIS RESULTS AND DISCUSSION

I.Project

In this project, calculations were made using 8 different bearing capacity methods. The results obtained from these are given in Table 6.

Table 6. Stress values obtained from bearing capacity methods

Bearing Capacity Methods	**Short q_t (kN/m^2)**	**q_t (kN/m^2)**
Terzaghi	1974.48	3033.06
Meyerhof	1926.30	2995.03
Hansen	1571.26	2240.23
Vesic	2002.14	2492.02
Prandtl	2161.54	4060.33
Bowles	292.34	292.34
Parry	642.86	642.86
Presiyometre	613.16	613.16

The soil bedding value of this project has been given as 6500kN/m^3. Based on this information, the stresses belonging to the foundation are given in Table 7.

Table 7. Foundation stress

	MC		**HS**	
Foundation Stress	$\sigma_{f(long)}$	$\sigma_{f(short)}$	$\sigma_{f(long)}$	$\sigma_{f(short)}$
kN/m^2	37.51	72.28	35.82	60.26

According to these tables, the highest q_t value was reached with the Prandtl bearing capacity method and the lowest q_t value with the Bowles bearing capacity method. In this project, the closest to the value obtained from the plaxis was reached by the Bowles bearing capacity method.

II.Project

In this project, calculations were made using 8 different bearing capacity methods. The results obtained from these are given in Table 8.

Table 8. Stress values obtained from bearing capacity methods

Bearing Capacity Methods	**Short Edge q_t (kN/m^2)**	**.ong Edge q_t (kN/m^2)**
Terzaghi	8434.57	13747.03
Meyerhof	29598.68	5285.10
Hansen	26110.49	39523.20
Vesic	28781.00	29914.24
Prandtl	14619.05	42124.67
Bowles	178.04	178.04
Parry	321.43	321.43
Presiyometre	916.46	916.46

The soil bedding value of this project has been given as 69000kN/m^3. Based on this information, the stresses belonging to the foundation are given in Table 9.

Table 9. Foundation stress

	MC		**HS**	
Foundation Stress	$\sigma_{f(long)}$	$\sigma_{f(short)}$	$\sigma_{f(long)}$	$\sigma_{f(short)}$
kN/m^2	93.84	176.64	39.01	77.28

According to these tables, the highest q_t value was reached with the Meyerhof bearing capacity method and the lowest q_t value with the Bowles bearing capacity method. In this project, the closest to the value obtained from the plaxis was reached by the Bowles bearing capacity method.

III.Project

In this project, calculations were made using 8 different bearing capacity methods. The results obtained from these are given in Table 10.

Table 10. Stress values obtained from bearing capacity methods

Bearing Capacity Methods	Short Edge q_t (kN/m^2)	Long Edge q_t (kN/m^2)
Terzaghi	2940.81	3704.95
Meyerhof	819.40	1007.89
Hansen	820.08	969.49
Vesic	992.63	1161.71
Prandtl	838.46	1495.20
Bowles	313.12	313.12
Parry	685.71	685.71
Presiyometre	570.29	570.29

The soil bedding value of this project has been given as 20000kN/m^3. Based on this information, the stresses belonging to the foundation are given in Table 11.

Table 11. Foundation stress

	MC		HS	
Foundation Stress	$\sigma_{f(long)}$	$\sigma_{f(short)}$	$\sigma_{f(long)}$	$\sigma_{f(short)}$
kN/m^2	458	725	173.4	287.8

According to these tables, the highest q_t value was reached with the Terzaghi bearing capacity method and the lowest q_t value with the Bowles bearing capacity method. In this project, the closest to the value obtained from the plaxis was reached by the Parry bearing capacity method.

IV.Project

In this project, calculations were made using 8 different bearing capacity methods. The results obtained from these are given in Table 12.

Table 12. Stress values obtained from bearing capacity methods

Bearing Capacity Methods	Short Edge q_t (kN/m^2)	Long Edge q_t (kN/m^2)
Terzaghi	3254.03	3566.54
Meyerhof	534.67	577.93
Hansen	538.34	570.33
Vesic	616.80	660.77
Prandtl	514.17	711.45
Bowles	184.92	184.92
Parry	407.14	407.14
Presiyometre	415.63	415.63

The soil bedding value of this project has been given as 24000kN/m^3. Based on this information, the stresses belonging to the foundation are given in Table 13.

Table 13. Foundation stress

	MC		HS
Foundation Stress	$\sigma_{f(long)}$	$\sigma_{f(short)}$	$\sigma_{f(long)}$
kN/m^2	3179.04	6372.96	1434.24

According to these tables, the highest q_t value was reached with the Terzaghi bearing capacity method and the lowest q_t value with the Bowles bearing capacity method. In this project, the closest to the value obtained from the plaxis was reached by the Terzaghi bearing capacity method.

V.Project

In this project, calculations were made using 8 different bearing capacity methods. The results obtained from these are given in Table 14.

Table 14. Stress values obtained from bearing capacity methods

Bearing Capacity Methods	q_t (kN/m^2)
Terzaghi	1882.01
Meyerhof	1069.31
Hansen	990.35
Vesic	1148.75
Prandtl	839.89
Bowles	205.73
Parry	428.57
Presiyometre	561.03

The soil bedding value of this project has been given as 20000kN/m^3. Based on this information, the stresses belonging to the foundation are given in Table 15.

Table 15. Foundation stress

	MC	HS-SS
Foundation Stress	σ_f	σ_f
kN/m^2	1575	312

According to these tables, the highest q_t value was reached with the Terzaghi bearing capacity method and the lowest q_t value with the Bowles bearing capacity method. In this project, the closest to the value obtained from the plaxis was reached by the Terzaghi bearing capacity method.

V. Conclusion

The following information can be obtained based on the results of the projects belonging to the five different soil classes we use:

- In the stresses calculated from the results obtained with the Plaxis program, it was seen that the highest stress was found with the Mohr-Coulomb soil model.
- Among the used bearing capacity methods, it has always been observed that the lowest value belongs to the Bowles bearing capacity theory.
- In the same soil layers, with the hardening soil model, lower stresses were obtained than the Mohr-Coulomb soil model.
- It has been observed that the Terzaghi bearing capacity method gives more suitable results in

cohesive soils.

REFERENCES

[1] Arş Gör Emrah Dağlı, Yrd Doç Dr Ömer Faruk Çapar and Yrd Doç Dr Ayşe Bengü Sünbül, "Analitik ve sayisal yöntemler ile kumlu zeminlerin taşima gücünün belirlenmesi," 2 International sustainable buildings symposium, 28-30 May 2015

[2] Brinkgreve, R B J (2002), "PLAXIS Finite Element Code for Soil and Rock Analysis- Version 8" Balkema, Rotterdam

[3] Castaldo, P , & De Iuliis, M (2014), "Optimal Integrated Seismic Design of Structural and Viscoelastic Bracing-Damper Systems", Earthquake Engineering & Structural Dynamics, 43(12), 1809-1827

[4] Duncan, J M and Chang, C Y (1970), "Nonlinear Analysis of Stress And Strain in Soils", Journal of the Soil Mechanics and Foundations Division, ASCE, Vol 96, pp 1629-1653

[5] Ercan, A , 2001 Deprem Yapıyı Nasıl Yıkar? Yer Araştırmalarında Zemin Emniyet Gerilmesi ve Yer İvmesinin Önemi Nedir? Kuzeybatı Anadolu Karesi Bölgesi Jeofizik Toplantısı, TMMOB Jeofizik Müh Odası, Balıkesir, 32-35

[6] Hansen, J B , 1961 The Bearing Capacity of Sand Tested by Loading Circular Plates 5 th International Conference on Soil Mechanic Foundation England:Vol 1, Paris

[7] Hsiung B B , Sakai, T (2016), "Failure Analysis of Underground Construction-Lessons Learned from Taiwan", Forensic Geotechnical Engineering, pp 197-208

[8] Keskin, M S , 2009 Güçlendirilmiş Kumlu Şevlere Oturan Yüzeysel Temellerin Deneysel ve Teorik Analizi Çukurova Üniversitesi, Fen Bilimleri Enstitüsü, Doktora Tezi,410s, Adana

[9] Meyerhof, G G , 1951 The ultimate bearing capacity of foundations, Geotechnique 2, No 4, pp 301-302

[10] Parry, R H G 1977, Estimating bearing capacity of sand from SPT values JGED ASCE 103:Gt 9 pp1014-1043

[11] Terzaghi, K , 1943 Theoretical Soil Mechanics John Wiley&Sons, New York

[12] Vesic, A S , 1973 Analysis of ultimate loads of shallow foundations, Journal of SM an FE Division, ASCE, vol 99, No SM1 pp 45-73

Analysis of Retaining Structures with Different Engineering Characteristics Used in Deep Excavations

M. KÜÇÜKBAYRAK[1] and H. F. PULAT[2]

[1] Katip Celebi University, Izmir/Turkey, kucukbayrakmehmet@gmail.com
[2] atip Celebi University, Izmir/Turkey, hfirat.pulat@ikc.edu.t

Abstract **- Nowadays, deep excavations are frequently used as a result of increasing urbanization and decreasing construction areas. In this paper, deep excavations and the design and selection criteria of these elements are discussed. Within the scope of the study, based on five different projects, 36 retaining structure system models with different features were created. The numerical analysis software PLAXIS 2D (version 2016.01), which is frequently used in the solution of geotechnical engineering problems and uses the finite elements method is used for the analysis of these models. With the 36 models created, the effects of vertical retaining element selection, material model selection and soil engineering properties on the performance of the retaining system were examined. As a result, it has been determined that while similar force and moment values are obtained in different vertical retaining elements, soil engineering properties and especially the choice of material model have serious effects on the results.**

Keywords **- bored pile, diaphragm wall, performance analysis, retaining structures, steel sheet pile.**

I. INTRODUCTION

etaining structures are building elements used to cut the soil surface in a vertical plane and create a ground surface on both sides at two different elevation levels [1]. Nowadays, the rapidly increasing population, especially in metropolitans, has caused intense construction, the decrease of the surface areas over time has made vertical construction above and below the ground necessary. This situation brings with it the obligation to carry out excavations in larger areas, which are getting deeper and deeper. In the face of all these needs, our civilization has developed retaining structures. Retaining structures can be built in excavation and filling areas, deep excavations, road bodies, as well as in different civil engineering sub-disciplines such as bridges, dams, ports, canals and flood prevention structures.

The concept of "Design of Retaining Structures" is among the subjects that constitute the first known and available written sources of engineering principles regarding soil mechanics. The emergence of rational methods for the design of retaining structures is based on the work of Gautier in 1717, who proposed a proportional method for gravity retaining structures. The studies of Coloumb (1776) and Rankine (1856), on the other hand, form the basis of our current understanding of lateral forces affecting the retaining structures [2]. During the design phase of retaining structures, soil parameters are used as expected. According to Genç [3]; Basic variables such as the internal friction angle of the soil, unit weight, cohesion value and the interaction between the retaining structure and the soil are used in the project phase. The lateral soil pressure distribution is obtained from the basic parameters such as lateral forces and shear forces that occur with the effect of these variables. As a result, the most economical type of retaining structure is designed, provided that it is a safe structure in accordance with engineering principles. The main purpose of this study is to select the most appropriate retaining element type, calculation and modeling method in terms of safety, aesthetics and financial aspects according to the project to be realized and the engineering properties of soil.

Within the scope of this study, a total of 36 system models were created based on five projects, and these models were analyzed by using the Plaxis 2D (version 2016.01) software. By systematically comparing the obtained analysis results, the effects of vertical retaining element selection, material model selection and soil engineering properties on retaining system design were interpreted.

II. METHODS

In this study, which is classified as flexible retaining structures and is now frequently used in deep excavations; diaphragm wall technique, bored pile wall coverings and sheet pile walls will be discussed. In addition, strut and anchor technique, which appear as internal braced members in deep excavation applications, will be discussed.

As it is known, after a deep excavation, retaining structures undertake the task of simulating the internal forces of the before excavation soil, which is in a state of balance. During the design of a retaining structure, the lateral earth pressures in active, passive and balance should be calculated by taking into account the internal forces of the soil and the forces caused by external effects [4].

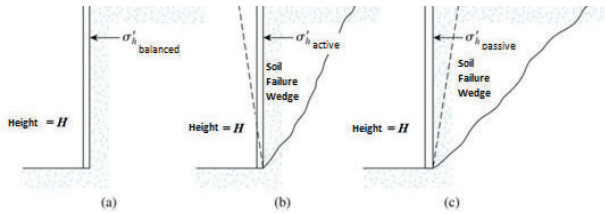


Figure 1: Earth pressures on a retaining structure of height *H*

Considering the soil and retaining structure model in Figure 1 (a), a retaining wall at H height, soil behind it, with γ unit weight, and a surcharge load q on the soil, are seen.

At this point, it will be useful to briefly mention about the Coulomb and Rankine Theories which are accepted as classical lateral earth pressure theories.

In order to simplify the solution of a problem related to lateral earth pressure, the assumptions to be made in Rankine Theory are listed below:

1. The soil is homogeneous, isotropic, cohesionless and has an internal friction angle,
2. The soil behind the retaining structure is in the form of a rigid wedge, triangular and failed along a plane,
3. The failure is handled as a 2-dimensional problem,
4. Friction resistance, evenly distributed throughout the failure surface,
5. Friction between the soil and structure will be neglected [5].

On the other hand, the most obvious and most important difference of Coulomb Theory stems from the assumptions of it. Coulomb Theory, contrary to Rankine Theory, accepts the existence of a friction angle between the soil and the structure. Coulomb interprets this as the angle of the combination of normal and shear forces on the wall, with the normal of the wall [2].

One of the important variables of the study is the material models of soil. There are different material models that can be used to solve engineering problems encountered in geotechnical engineering. Significant differences can arise between the results of the analyzes performed with these models. In the analyzes made in this study, Mohr-Coulomb model and another material model, which is suitable for all five projects and their soil type, was used. Afterwards, the analysis results were compared and the differences were tried to be revealed.

The Mohr-Coulomb model is a simple and well-known linear elastic-perfectly plastic behavior model. This model is a model that can be used as the first approach in modeling the soil behavior, thus contributing to a general idea about the problem.

In the Hardening Soil model, irreversible deformations due to primary loading is modeled by shear hardening. Compression hardening is operated to model plastic deformations which are irreversible caused by primary compression in oedometric loading and isotropic loading. In other words, there are both types of hardening in this model. This model;

- uses the theory of plasticity rather than the theory of elasticity,
- includes soil dilatation and
- takes into account the yielding behavior [6].

As is known, Plaxis 2D (version 2016.01) software uses finite element method in its analysis and calculations. As in every area also in geotechnical engineering, the finite element method aims to simplify a complex solution by dividing a big problem, into smaller problems. Since it would be difficult and time-consuming to analyze a structure with complex geometry, it is based on the logic of creating a finite element mesh by dividing the structure into small parts and reaching a conclusion about the whole from the part [7].

III. DETAILS OF PROJECTS

This study includes the analysis of fictive projects prepared using five different geotechnical projects. Based on these five projects, a total of 36 different models were created. The general features of these five projects are as in Table 1.

Table 1: General Features of the Projects

	Project 1	Project 2	Project 3	Project 4	Project 5
Soil Types	Mixed	Cohesive – Non Cohesive	Cohesive – Non Cohesive	Cohesive – Non Cohesive	Cohesive – Non Cohesive
Retain Struct Types	Dia Wall – Bored Pile	Dia Wall – Bored Pile	Dia Wall – Bored Pile	Dia Wall – Bored Pile	Sheet Pile – Dia Wall
Excav Depth	7m	4m	4 5m	9m	4 5m
İntern Braces	Double strut	Single strut	Single Anhor	Double Anhor	-
Embed Length	13 m	12 m	6 5 m	5 5 m	6 5 m
Cant Length	7 m	4 m	4 5 m	7 m	4 5 m
Total Length	20 m	16 m	11 m	12 5 m	11 m

In the study, firstly each project created was modeled, and then three variables (vertical retaining element type, soil material model and soil formation) were changed respectively and the models that emerged with different combinations were analyzed. While carrying out all these studies, issues such as the consistency of the project data and the acceptability of deformations in the system as a result of the analysis were excluded from the scope of the study. The aim is to evaluate the differences in results between models established with exactly the same features except for variables.

Since a mixed soil profile with granular and plastic material layers in the same profile was created in Project 1, an alternative soil profile was not produced. However, the soil profile was formed from completely plastic material layers in Project 2, and completely granular material layers in Project 3, Project 4 and Project 5. For this reason, a soil profile with the opposite feature of these soil profiles was also approximately created and the analyzes were repeated in order to be an alternative. Thus, the effect of the soil formation being granular or plastic on the performance of the retaining systems has also been observed.

All of the system modeling and analysis were carried out with the Plaxis 2D (version 2016.01) software. While creating the models, some assumptions have been made when determining alternative vertical retaining elements, determining the parameters of the opposing soil formation or making transition between material models.

- The variables “Model”, “Elements”, “Length”, “Force”, and “Time” were chosen as “Plain strain”, “15-Noded”, “m”, “kN”, and “day”, respectively.
- Struts were modeled using the "node-to-node anchor" element.
- For anchors, "anchor" element for free length and "geogrid" element for bond length were preferred.

- Finite element meshes were created by selecting the "medium" option for level of global coarseness. No specific level of coarseness has been defined for any element in any project.
- Thermal parameters were not taken into account.
- For each construction phase in the "Staged construction" section, the periods to be deemed reasonable according to the relevant construction phase have been selected and these periods have been entered in the "Time interval" section.
- In the required construction phases, groundwater flow calculation was also defined in the software and this calculation was made.
- Material properties of C30/37 concrete were used for all reinforced concrete elements and the modulus of elasticity (E) was included in the calculations as 32000 MPa [8].
- Poisson ratio of concrete was chosen to be 0.15 [5].
- The modulus of elasticity (E) of steel was included in the calculations as 210000 MPa [9].
- Poisson ratio of steel was chosen to be 0.30 [5].
- Surcharge loads defined in the system for surrounding buildings were determined as 15 kN/m² per floor in order to stay on the safe side [9].
- Surcharge loads defined in the system for the surrounding roads were determined as 15 kN/m² for secondary roads in order to stay on the safe side [9].
- For retaining structure systems, long-term parameters were used in the analyzes.
- While transitioning between bored pile and diaphragm wall elements, for 1 meter length into the sheet plane, the alternative element size was determined so that both elements have the same area value.
- When transitioning between diaphragm wall and steel sheet pile elements, since the material structures and cross-sectional areas are very different from each other, the dimensions of the new element are determined so that the bending stiffness ($E \times I$) values of the materials are equal.
- While creating soil profiles which are the opposite of the first profiles, the elasticity module, Poisson ratio, cohesion, internal friction angle, dilatation angle and permeability coefficient values of the soils were changed.
- While determining these values for the new profiles to be created, the studies of researchers, which are widely accepted in the geotechnical engineering literature, have been used.
- Within the ranges given for the relevant soil type, the relevant value has been chosen to be at approximately the same position within the range given for the new soil profile.
- A parameter value for a soil type lies at a certain position within a range. The new value of that parameter for the opposing soil profile was chosen to be at approximately the same position within its range.
- The selected data were chosen to reflect the characteristic features of the soil type represented.
- While material sets were created with different soil material models, soil elasticity modulus (E) and Poisson ratio (v) were used in Mohr-Coulomb model. In the hardening soil model, it was necessary to work with the soil oedometric deformation module (E_{oed}), triaxial loading deformation module (E_{50}^{ref}) and triaxial unloading deformation module (E_{ur}^{ref}).
- The following equations have been used for E_{oed}, E_{50}, and E_{ur}.

$$E_{oed} = E_{50} \quad (1)$$

$$E_{ur} = 3E_{50} \quad (2)$$

IV. Analyses, Results, and Discussions

Based on a total of five projects, analyzes were carried out on 36 system models created. The results obtained from the PLAXIS 2D (version 2016.01) software have been examined in detail.

Table 2: Analysis Results - 1

Project	Soil Profile	Mater Model	Reta Wall	Msf	u	u_x	u_y
Proj 1	Mixed	MC	D W	3 15	27 03	6 40	26 76
			B P	3 14	26 79	5 73	26 55
		HS	D W	3 14	8 42	4 22	8 10
			B P	3 14	7 91	3 47	7 72
Proj 2	Sandy Soil	MC	D W	3 10	16 87	2 50	16 86
			B P	3 07	17 00	2 49	17 00
		HS	D W	3 10	4 98	1 64	4 96
			B P	3 08	4 75	1 25	4 74
	Clayey Soil	MC	D W	2 80	19 88	3 51	19 88
			B P	2 77	20 02	3 50	20 02
		HS	D W	2 82	4 56	1 91	4 52
			B P	2 77	4 46	1 64	4 44
Proj 3	Sandy Soil	MC	D W	2 17	8 06	2 63	8 03
			B P	1 99	8 14	2 68	8 12
		HS	D W	2 17	3 70	1 36	3 69
			B P	1 99	3 70	1 38	3 70
	Clayey Soil	MC	D W	1 72	7 96	5 40	7 39
			B P	1 58	8 24	5 54	7 61
		HS	D W	1 72	2 59	1 09	2 58
			B P	1 58	2 61	1 11	2 61
Proj 4	Sandy Soil	MC	D W	1 76	13 37	10 27	11 89
			B P	1 72	13 52	10 36	11 95
		HS	D W	1 76	8 03	6 57	5 09
			B P	1 72	7 86	6 67	5 02
	Clayey Soil	MC	D W	1 53	18 88	15 72	13 62
			B P	1 50	19 09	16 12	13 75
		HS	D W	1 53	10 57	8 68	5 79
			B P	1 50	10 30	8 74	5 64
Proj 5	Sandy Soil	MC	S P	1 58	10 31	1 85	10 31
			D W	1 59	10 31	2 08	10 31
		HS	S P	1 60	4 02	2 08	4 02
			D W	1 62	4 02	2 37	4 02
	Clayey Soil	MC	S P	1 53	10 00	5 01	9 81
			D W	1 51	10 04	5 37	9 82
		HS	S P	1 53	2 79	1 92	2 79
			D W	1 52	2 79	2 15	2 79

Table 3: Analysis Results - 2

Project	Soil Profile	Mater Model	Reta Wall	P_{u_x}	P_{u_y}	P_s	P_m
Proj 1	Mixed	MC	D W	6 38	14 77	198 20	357 90
			B P	5 72	13 33	223 60	503 30
		HS	D W	4 22	3 33	192 30	297 00
			B P	3 48	3 09	208 50	409 80
Proj 2	Sandy Soil	MC	D W	2 40	11 61	73 16	102 90
			B P	2 08	10 42	73 77	95 95
		HS	D W	1 72	2 92	67 87	107 80
			B P	1 35	2 78	67 11	113 00
	Clayey Soil	MC	D W	2 91	13 97	60 66	68 95
			B P	2 62	12 68	58 23	-83 86
		HS	D W	1 79	2 62	-72 06	-104 90
			B P	1 53	2 42	-78 21	-114 40
Proj 3	Sandy Soil	MC	D W	2 44	-2 01	145 80	166 00
			B P	2 44	-3 05	142 80	146 90
		HS	D W	0 89	-1 53	152 90	168 80
			B P	0 86	-2 10	150 60	155 10
	Clayey Soil	MC	D W	5 54	-3 53	160 30	204 00
			B P	5 67	-4 84	156 80	185 20
		HS	D W	0 66	-1 40	158 10	172 90
			B P	0 65	-1 97	156 70	166 40
Proj 4	Sandy Soil	MC	D W	10 31	-1 68	98 83	184 10
			B P	10 38	-2 21	108 40	231 20
		HS	D W	6 61	-2 31	110 30	222 10
			B P	6 80	-2 63	125 20	297 60
	Clayey Soil	MC	D W	15 78	-4 31	111 50	227 00
			B P	16 21	-5 17	122 10	276 50
		HS	D W	8 73	-3 52	128 90	272 80
			B P	8 90	-3 99	147 70	368 30
Proj 5	Sandy Soil	MC	S P	1 89	2 99	-37 72	-93 21
			D W	2 15	2 16	-37 71	-90 25
		HS	S P	2 20	1 12	-45 29	-120 30
			D W	2 50	0 59	-46 52	-124 90
	Clayey Soil	MC	S P	5 05	2 00	-20 87	-26 63
			D W	5 57	0 93	-23 07	-29 25
		HS	S P	2 06	0 69	-39 01	-90 41
			D W	2 36	0 31	-40 77	-95 25

Effects of Vertical Retaining Element Selection on Results

As the method followed in the study, while choosing the equivalent element size between bored piles and diaphragm wall elements, attention has been paid to ensure that the cross-sectional area of these elements per 1 m perpendicular to the page plane is the same.

At this point, it was seen that the axial stiffness ($E \times A$) values calculated for bored pile and diaphragm wall elements were almost equal for all models, since the selected concrete classes were also the same.

In addition, when the bending stiffness values ($E \times I$) are considered, in bored piles and diaphragm wall elements with the same "cross-sectional area / m" values, it has been determined that the moment of inertia of bored pile elements and therefore the bending stiffness values are equal to 1.5 to 2 times these values of the diaphragm wall elements.

It was also mentioned that the dimensions of the new element were determined so that the bending stiffness ($E \times I$) values of the materials were equal when transitioning between the diaphragm wall and steel sheet pile elements. In this case, it has been determined that the axial stiffness ($E \times A$) of the diaphragm wall in the present example is equal to 2 times that of the steel sheet pile.

Although there is a 1-fold difference between the bending stiffness of the diaphragm walls and the bored piles, and another 1-fold difference between the axial stiffness of the steel sheet piles and the diaphragm walls, this did not result in significant differences between the displacements of the system and the displacement and forces of the plate elements.

As mentioned, although there were not significant differences in the results, some differences were detected at small ranges. The most consistent difference among these has emerged in the safety factors of the systems. Although the difference is very small, it has been determined that in projects modeled using both diaphragm wall and bored pile alternatives, the system safety factors decrease slightly in bored pile systems. In the comparison of steel sheet pile - diaphragm wall, it was observed that there was a very small decrease in the safety factors of steel sheet pile systems.

In projects modeled with bored pile - diaphragm wall elements alternatives (Project 2, Project 3 and Project 4), it has been determined that the maximum displacements in the x direction of the system are slightly higher in bored piles. In Project 5, which has a cantilever retaining system and modeled with a steel sheet pile - diaphragm wall, it is seen that the maximum displacements in the x direction of the diaphragm walls are significantly higher.

Effects of Soil Material Model Selection on Results

When analyzing the analysis results of the systems analyzed by choosing different soil material models, it is seen that there are very significant and consistent differences between different material models, especially in terms of displacement values. In almost all of the examples available, when the Mohr-Coulomb (MC) material model was preferred, the displacements increased significantly compared to the Hardened Soil (HS) model. It is seen that these increases occur between 0.5 and 4 times in total system displacements and between 0.5 and 4 times in plate element displacements. Considering that this is due to the use of different modulus of elasticity between material models and the fact that the unloading stiffness is considered to be significantly increased, it can be easily said that the results are consistent.

Considering the system and plate displacements, it was not meaningful that only the settlement values of the Project 1 models and the three displacement values ($u_{x_{max}}$, $u_{y_{min}}$ (settlement), $P_{u_{x_{max}}}$) of the Project 5 "sand soil" models were higher in the Hardening Soil model compared to the Mohr-Coulomb model. When the analyzes were rigorously repeated, the results did not change. Although it is striking as a result of the investigations that the soil layers of this model set have the highest internal friction angle values among all models, it is thought that it may be useful to consider the reasons behind these results in a different study.

When the section forces occurring in the plate elements are

examined, the axial force values have increased for all models when the Hardening Soil (HS) material model is preferred. In addition, it has been observed that the shear force and moment values increase in most of the models when Hardening Soil (HS) material models are preferred.

As an interesting result, it was found that the choice of soil material model has almost no effect on the overall safety factors of the systems.

Performance of Retaining Systems in Different Soil Formations

By leaving all the remaining parameters the same, performance differences in retaining systems, which occur only by changing the soil formations of the systems, have been examined. It is seen that there are unstable findings about the performances of retaining systems in different soil formations.

As the most important and stable finding, it is observed that the safety factors obtained as a result of the analysis of exactly the same retaining systems on granular (sand) and plastic (clay) soils, decreased significantly for each model group in plastic soils.

Considering the force and moment values occurring in the cross sections of plate elements, it is seen that the force and moment values occurring in plastic soils are steadily higher.

V. Conclusions

As a result, when the results obtained from the study are examined, it is seen that the selection of the appropriate material model has a very serious effect on the force and displacement values of the system and elements.

On the other hand, in cohesive soils, the force and displacement values affecting the retaining elements increased due to the low soil shear strength effect as expected. As a result, it can be easily said that the system safety factors will decrease significantly even if larger sections are not selected.

It has been observed that there were no major differences in force and displacement values between different retaining elements. Here, it can be stated that the fact that all of the retaining elements used are fragile retaining elements and that the changes of the retaining element do not cause a change in the active-passive force balance of the system are effective on this result.

References

[1] Clayton CWRBAMJ Earth-pressure and earth-retaining structures, 3rd ed : CRC Press; 2013
[2] D Coduto, Foundation Design: Principles and Practices: Prentice-Hall; 2001
[3] D Genç, Zemin Mekaniği ve Temeller: TMMOB Jeoloji Mühendisleri Odası; 2011
[4] K Özaydın, Zemin Mekaniği: Birsen Yayınevi; 2016
[5] *J. Bowles, Foundation Analysis and Design, 5th ed.: McGraw-Hill; 1997*
[6] PLAXIS PLAXIS Material Models Manual; 2019
[7] S Singiresu, The Finite Element Method in Engineering, 4th ed : Elsevier Science & Technology Books; 2004
[8] TSE Betonarme Yapıların Tasarım ve Yapım Kuralları 2000 TS 500
[9] Look B Handbook of Geotechnical Investigation and Design Tables: Taylor & Francis; 2007

A Novel Logarithmic Amplifier with Second Generation Current Conveyors

M. ISTANBULLU[1] and M. AVCI[1]

[1] Çukurova University, Adana/Turkey, mm.istanbullu@gmail.com
[1]Çukurova University, Adana/Turkey, mutluavci@gmail.com

***Abstract* – This study presents the design and the complete analysis of a novel logarithmic amplifier using second-generation current conveyors. The circuit utilizes exponential current-voltage characteristics of a pn-diode to obtain the logarithm of the input signal. Unlike the opamp-based logarithmic amplifiers, the proposed design does not suffer from instability since the active component in the feedback loop. The simulation results show that the logarithmic behaviour of the design could be used in various applications.**

***Keywords* – Logarithmic amplifier, current conveyors, pn-diode.**

I. Introduction

A Logarithmic amplifier, or logger, is an amplifier that yields an output proportional to the logarithm of the input voltage. Logarithmic amplifiers can be used in the circuits to convert a high dynamic range input into an output with a determined range [1-2]. In recent years, logger circuits are designed to achieve low power, low noise for neural and biopotential signal acquisition in which input signals vary between several microvolts to millivolts [3-8].

In the literature, various circuit topologies to realize a logarithmic amplifier have been introduced. Generally, pn diode, light-emitting diode, bipolar junction transistor, and MOSFETs in weak inversion are utilized to obtain logarithmic characteristics. However, trans-impedance-based topologies suffer from several problems such as static gain, limited output dynamic range, temperature compensation, and necessity of voltage to current converter for current mode operation [9]. The circuit presented here is based on the second-generation current conveyor that is working on current mode operation.

The second-generation current conveyor (CCII) offered by Sedra and Smith [10] is commonly used in filters and oscillators [11]. Current sourcing (CCII+), as well as sinking (CCII-) current conveyors, are available in the literature.

In this work, a novel logarithmic amplifier is proposed using two series CCII+ conveyors. A pn-diode is utilized to obtain logarithmic characteristics. Circuit design and simulations are performed on the PSPICE environment.

II. Second Generation Current Conveyor (CCII+)

The block diagram of the second-generation current conveyor is given in Figure 1. Basically, CCII can be considered as a primary block of the filter circuits. Additionally, they are used in digital to analog converters, amplifiers, and oscillators. CCIIs are suitable for low voltage, high frequency, and low power applications. Y port accepts the input voltage and transfers it to the X port. Then, the input current at the X port is conveyed to the Z port. The signs in the CCII+ and CCII- blocks express the same amount of output current but in reverse directions.

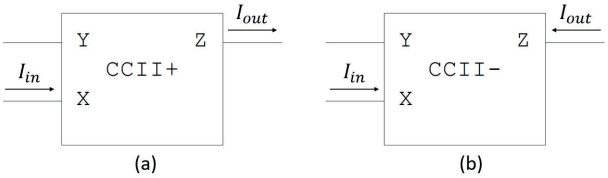


Figure 1. Block diagrams of the second generation (a) current source, (b) current sink current conveyors

The complete circuit schematic of the CCII+ block shown in Figure 1(a) is given in Figure 2. General characteristics of the second generation current conveyor can be obtained by direct circuit analysis. As shown in the figure, the Y terminal is the non-inverting input of the operational amplifier, which has high input impedance. Thus,

$$I_Y = 0 \tag{1}$$

Voltages on the non-inverting and inverting terminals are approximately the same due to the virtual ground of the opamp.

$$V_X = V_Y \tag{2}$$

In order to obtain the output current definition, assume I_X current flows from the X terminal towards the ground.

$$I_X = \frac{V_1 - V_2}{R} \tag{3}$$

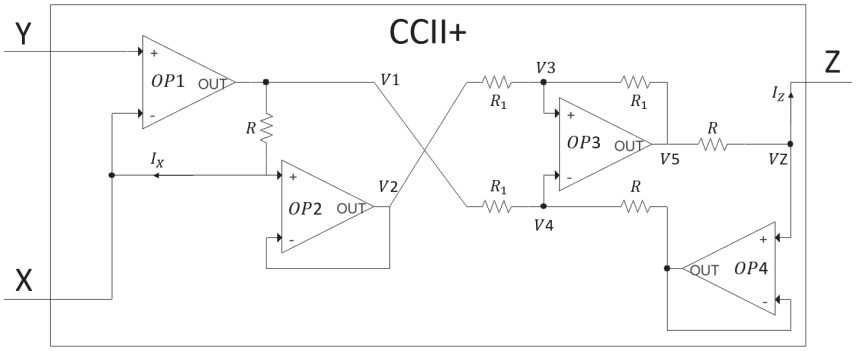


Figure 2. The complete circuit of the second generation current conveyor

$$I_1 = \frac{V_2 - V_3}{R_1} = \frac{V_3 - V_5}{R_1} \tag{4}$$

Then,

$$V_5 = 2V_3 - V_2 \tag{5}$$

and similarly,

$$V_Z = 2V_4 - V_1 \tag{6}$$

As shown in Figure 2, the voltage of the inverting terminal is equal to that of the non-inverting terminal, hence,

$$V_4 = V_3 \tag{7}$$

by using (5) and (6),

$$V_5 - V_Z = V_1 - V_2 \tag{8}$$

similarly,

$$\frac{V_5 - V_Z}{R} = \frac{V_1 - V_2}{R} \tag{9}$$

Therefore, the output current is equal to the input current.

$$I_Z = I_X \tag{10}$$

Hence, the current conveyor given in Figure 2 operates as the current-sourcing mode.

III. Logarithmic Amplifier Circuit

To achieve the proposed logarithmic amplifier, the circuit given in Figure 3 is considered. In the circuit, the input is a voltage on the Y terminal. Using the properties of CCII+, it can be observed that the same current also flows through the D1 diode. The input current of the first CCII+ is as follows.

$$I_{X_1} = \frac{V_{in}}{R_1} \tag{11}$$

where V_{in} is the input voltage. The current that conveyed to the output of the first CCII,

$$I_D = I_S e^{(V_D / nV_T)} \tag{12}$$

where, $I_D = I_{Z_1} = (V_{in} / R_1)$. Thus,

$$\frac{V_{in}}{R_1} = I_S e^{(V_D / nV_T)} \tag{13}$$

Here, I_S is the reverse saturation current, and V_T stands for the thermal voltage.

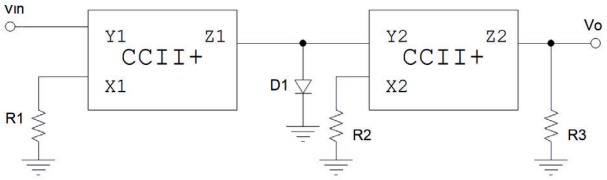


Figure 3. Proposed logarithmic amplifier circuit

The input voltage of the second CCII+ can be obtained by rearranging (12) as following.

$$V_D = nV_T \ln\left(\frac{V_{in}}{R_1 I_S}\right) \tag{14}$$

where V_D is the voltage appears on the D_1 doide ($V_{Y_2} = V_D$). The voltage on the X_2 terminal is the same as the voltage on Y_2 ($V_{X_2} = V_D$). Then, the input current at the X_2 terminal and conveyed to the Z_2 output are as follows.

$$I_{Z_2} = I_{X_2} = \frac{V_D}{R_2} \tag{15}$$

Finally, the output voltage at the Z_2 terminal of the second CCII+ can be obtained as,

$$V_0 = \frac{R_3}{R_2} n V_T \ln\left(\frac{V_{in}}{R_1 I_S}\right) \quad (16)$$

Therefore, the output voltage is a function of logarithm of the input signal, as obvious in (15). The control of the amplifier gain can be altered by changing the R_2, R_3 resistor values.

IV. Simulation Results And Discussion

The proposed circuit is simulated is PSPICE environment. The both CCII+ circuits were designed by using TL081 opamp. 1N4148 diode model was obtained from PSPICE library. The ideality factor n and reverse saturation current I_S for the diode are 1.836 and 2.628nA, respectively.

Opamps in the proposed circuit are biased symmetrically with $\pm 5V$ power supply. The input DC voltage V_{in} is swept from 0V to 5V. The simulation result is plotted in Figure 4. The output voltage is proportional to the input voltage, as shown in the figure.

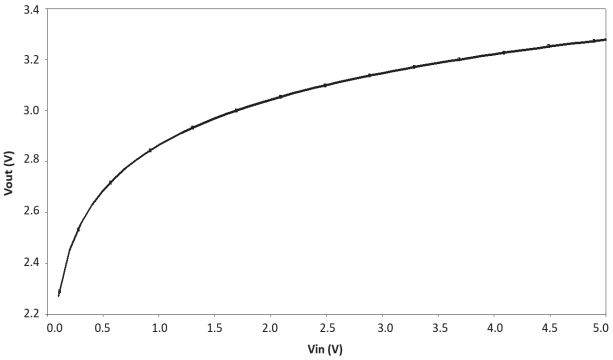


Figure 4. The output of the proposed logarithmic amplifier

To test the feasibility of the proposed study, the gain of the design for three load resistances (5kΩ, 3kΩ, 1kΩ) are also presented. Figure 5 shows the gain in decibel to different frequency values.

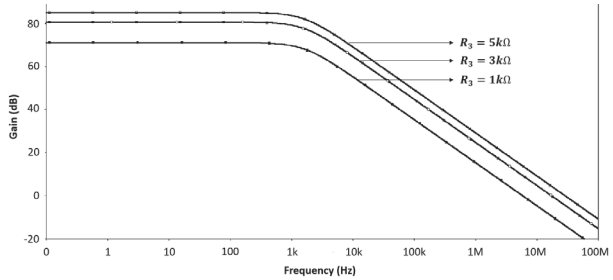


Figure 5. Frequency response of the proposed logarithmic amplifier

In this paper, a novel logarithmic amplifier is proposed using two second generation current conveyor. Notice that the gain is simply adjustable by changing external resistors. The operation of the proposed circuit is suitable not only for voltage inputs but also current. The circuit presented here can be used for different purposes, such as analog signal compression, biomedical signal processing, or logarithmic ratio computation in various instrumentation units.

References

[1] K Hong-Sun, M Ismail, and H Olsson, "CMOS limiters with RSSIs for Bluetooth receivers," in *Proceedings of the 44th IEEE 2001 Midwest Symposium on Circuits and Systems. MWSCAS 2001*, 2001

[2] Z Yong, C Lei, Z Xiao-Jun, and L Zong-Sheng, "A parallelamplification parallel-summation logarithmic amplifier for UHF RFID reader," in *2009 IEEE 8th International Conference on ASIC*, 2009

[3] J Ramos, J Ausin, J Duque-Carrillo, and G Torelli, "Design of limiting/logarithmic amplifier for wideband bioimpedance measuring devices," in *2010 Biomedical Circuits and Systems Conference (BioCAS)*, 2010

[4] Z Derafshi and J Frounchi, "Low-noise low-power front-end logarithmic amplifier for neural recording system," *Int. J. Circuit Theory Appl.*, vol 42, no 5, pp 437–451, 2014

[5] Y Sundarasaradula and A Thanachayanont, "A Low-Noise, LowPower, Wide Dynamic Range Logarithmic Amplifier for Biomedical Applications," *J. Circuits, Syst. Comput.*, vol 27, no 7, 2018

[6] E Bahrami and H Shamsi, "Ultra-low-power front-end CMOS true logarithmic amplifier for biopotential signal acquisition applications," in *23rd Iranian Conference on Electrical Engineering*, 2015

[7] R Chebli and M Sawan, "Chopped Logarithmic Programmable Gain Amplifier intended to EEG acquisition interface," in *25th International Conference on Microelectronics (ICM)*, 2013

[8] Y Zili and M Gerard, "A programmable time-gain compensation (TGC) amplifier for medical ultrasonic echo signal processing," in *IEEE International Conference on Solid-State and Integrated Circuit Technology (ICSICT)*, 2008

[9] M Gu and S Chakrabartty, "Design of a Programmable Gain, Temperature Compensated Current-Input Current-Output CMOS Logarithmic Amplifier," *IEEE Trans. Bio. Med. Cir. Syst.*, vol 8, no 3, pp 423–431, 2014

[10] A Sedra and K C Smith, "A second generation current conveyor and its applications," *IEEE Trans. Circuit Theory*, vol 17, pp 132–134, 1970

[11] I A Khan, H Jaidi, and H Mehmood, "Multifunctional translinear-C-current mode filter," *Int. J. Electron.*, vol 87, no 9, pp 1047–1051, 2000

A Rail to Rail, Low Voltage Operational Amplifier in 0.18µm CMOS Technology

M. ISTANBULLU[1]

[1] Çukurova University, Adana/Turkey, mm.istanbullu@gmail.com

Abstract **– A two-stage opamp operating at 0.9 V and consuming 5 µA has been proposed in a TSMC 0.18 µm CMOS technology. Rail to rail operation at the input stage is achieved by complementary transistor pairs with transconductance control. The rail to rail output stage is designed in class AB. All design and simulations are performed on the Cadence environment. The simulated gain of the opamp is 67 dB, and the unity gain frequency is 1.7 MHz with a 3 pF, 100 kΩ parallel load.**

Keywords **– Opamp, rail-to-rail, TSMC 0.18 µm technology**

I. Introduction

DEVICE feature scaling is a critical issue to reduce supply voltages, currents and thus, power consumption of the device [1]. However, reducing the required supply voltage may lead to a decrease in the dynamic range of the device [2]. For a low-voltage opamp, extending the operation rail-to-rail can increase the dynamic range. A further problem with the low-voltage opamp is the requirement of the additional circuit blocks since the threshold voltage of transistors is not scaled with the supply voltage [3].

In the literature, several designs are presented for low-voltage, power efficient opamps [4][5][6][7]. These studies have rail-to-rail input stages and current efficient output stages. The two gate to source voltage plus two saturation voltages of four transistors in the input stage determine the minimum supply voltage, approximately 2.5V [5].

In the literature, input level shifting [8], current-driven bulk [9], floating-gate transistor [10], bulk driving [11][12], and DT-MOS [13] techniques are reported in order to operate opamp containing high threshold voltage transistors, close to 1V supply voltage. Although these techniques overcome the threshold problem, they have some primary disadvantages. In the level shifting method, the usage of resistors increases both noise and device size. A less gain-bandwidth (GBW) handicap arising from the smaller transconductance is available in the bulk-driven as well as floating gate methods. Producing DT-MOS and floating gate transistors require expensive and non-standardized processing steps.

In this work, a rail-to-rail opamp operating at sub-1V supply voltage is presented. The p-channel and n-channel transistors have -.05V and 0.45V threshold voltages and biased in the subthreshold region. The circuit topology of the amplifier given in Figure 1 is described in the following subsections.

II. Circuit Topology

A. Rail-to-Rail Input Stage

The rail-to-rail operation is achieved by using two differential pairs (n-channel and p-channel) in a parallel manner. P-channel M4 and M7 transistor pair are ON for low common-mode voltages, whereas n-channel M5 and M6 transistor pair are ON for high common-mode voltages. The minimum value of the required power supply is obtained by the sum of two gate to source voltage and two saturation voltages as given in (1).

$$V_{PS\_\min} = V_{GS\_n} + V_{GS\_p} + V_{DS\_satn} + V_{DS\_satp} \qquad (0.1)$$

where V_{GS_n} and V_{GS_p} stand for gate to source voltages of n- and p-channel input transistors, whereas V_{DS_satn} and V_{DS_satp} are saturation voltages of the M18 and M8 transistors (current source), respectively. For 0.9V power supply, the input transistors are biased in the subthreshold region $(2x0.3V + 2x0.15V = 0.9V)$.

The variation of transconductance value over the input common voltage range is the primary problem in the presence of the complementary input stage. In the common voltage range midpoint, both transistors in the input stage are ON state at the same time; therefore, the total drain to source current is twofold of that of the current when one of the input pair is in ON state. Thus, an additional circuit block is required to keep the transconductance value constant. In the presented circuit topology, M2 and M1 transistors, as the current switch, control the transconductance. In this way, the input stage provides a constant output current to the summing block. The summing circuit block contains a current mirror (M12 and M10) and a common gate stage (M13 and M16) and requires one gate to source strong inversion voltage plus two drain to source voltages for proper operation.

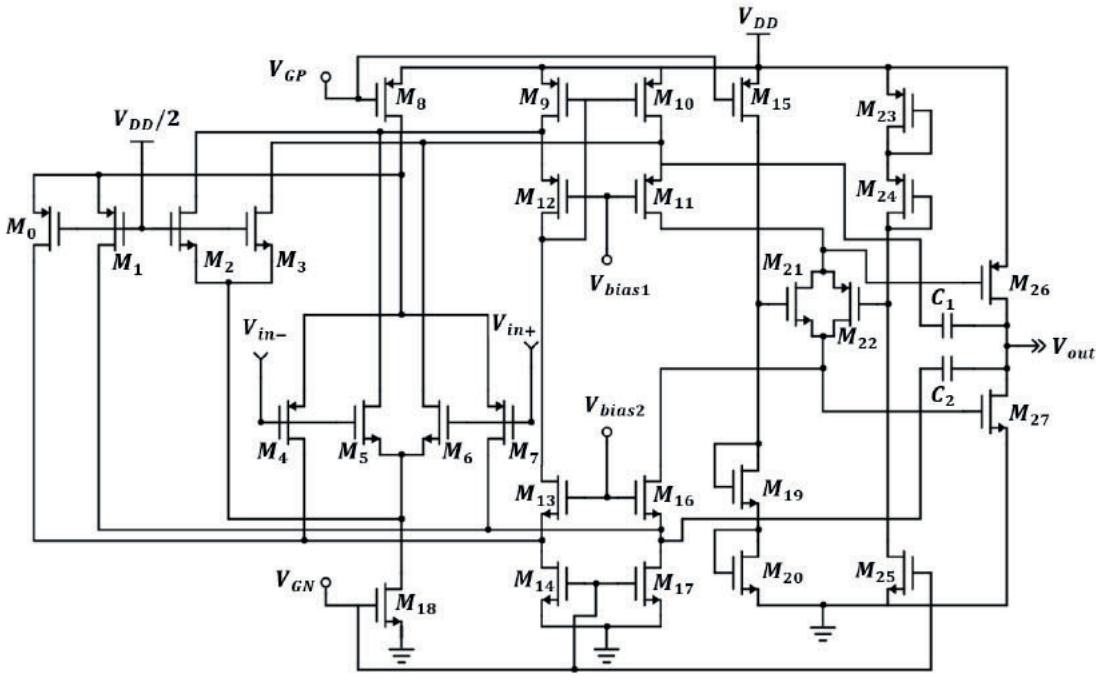


Figure 1. The complete circuit topology of the proposed opamp

A. *Class AB Push-Pull Output Stage*

The M27 and M26 output transistors are designed in a common source configuration and biased in class AB type to guarantee rail-to-rail operation. In the circuit presented here, the control of the class AB operation is done by M21 and M22 transistors. These control transistors are driven by the currents of M16 and M11 transistors. The gate terminals of the control transistors are supplied by a constant voltage from two diode-connected M19-M20 and M24-M23 transistors. The bias current of the output transistors is determined by the diode-connected transistors (M19, M20, M24, M23), control transistors (M21-M22), and output transistors (M27, M26). In this configuration, since the gate to source voltages and device dimensions of the M19 and M21 transistors are the same, M20 and M27 transistors have the same gate to source voltages. Hence, the quiescent current of the output transistors can be obtained as the ratio of the aspect ratios of M27 and M20 transistors. It is apparent from Figure 1 that, the connected diode transistors require two gate to source voltages and one drain to source voltage for proper operation. In the circuit presented here, current source transistors are biased in the saturation region, whereas the other output transistors are biased in the sub-threshold region.

B. *Device Dimensions*

Dimensions of all the transistors in the topology are given in Table 1. The frequency compensation of the opamp is done by cascoded Miller frequency technique since the amplifier is designed for low current consumption.

Table 1: Transistor dimensions and capacitor values.

Transistor Labels	W(µm)/ L(µm)
M5, M6, M2, M3	50/0.36
M4, M7, M0, M1	165/0.36
M13, M16	10/0.36
M12, M11	33/0.36
M18, M14, M17	10/6
M8, M9, M10	33/6
M25	2.5/6
M15	8.25/6
M21, M19	30/0.18
M22, M24	99/0.18
M20	5/0.18
M23	16.5/0.18
M27	60/0.18
M26	198/0.18
Capacitor Labels	**Value**
C1, C2	1 pF

III. Simulation Results And Conclusion

The simulation results of the proposed CMOS amplifier agree with the design specifications. The gain-bandwidth is obtained as 1.7 MHz for a 3 pF load with a phase margin of 60° using a 1 pF Miller capacitor. The DC gain is observed as 67 dB. The simulated gain and phase margin characteristics are given in Figure 2 when 0.9 V power supply and 0.4 V common-mode input voltage. The current consumption is 5 µA with 0.9 V supply voltage.

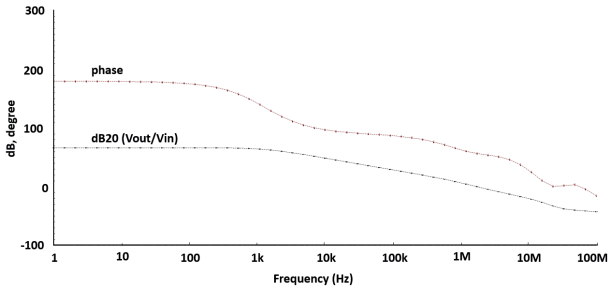


Figure 2. Gain and phase margin of the opamp

300 mV (large) and 50 mV (small) input step voltages in 250 kHz frequency are applied to the opamp in order to obtain slew rate value. Simulation results when the proposed opamp configured in a unity-gain buffer configuration are given in Figure 3. The simulated slew rate is 0.6 V/µs.

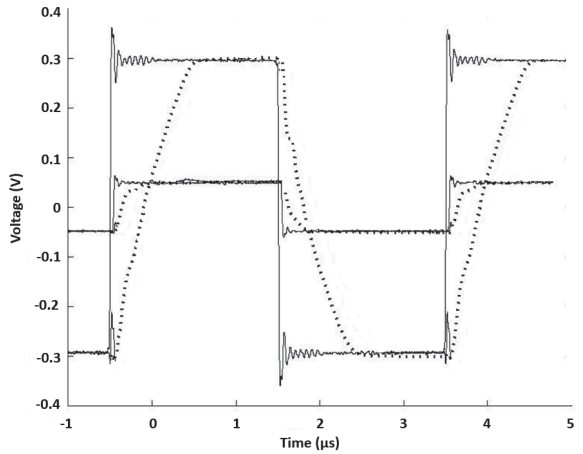


Figure 3. Simulated input (line) and output (dashed line) signal for buffer configuration

It is observed that, when the opamp resistively loaded with a 1 kΩ in the unity gain configuration, the output signal is clipped at V_{DD}-16mV and V_{SS}+10mV.

The original amplifier design is simulated on Cadence EDA software with 0.18µm TSMC process device technology. The opamp presented here operates in low voltage and low power conditions. The amplifier shows rail to rail operation and high gain-bandwidth characteristics.

REFERENCES

[1] R J Baker, *CMOS: Circuit Design, Layout, and Simulation* Wiley, 2019

[2] C J Huang and H Y Huang, "A low-voltage CMOS rail-to-rail operational amplifier using double p-channel differential input pairs," *Proc. - IEEE Int. Symp. Circuits Syst.*, vol 1, pp I–673, 2004

[3] P E Allen and D R Holberg, *CMOS Analog Circuit Design* Oxford University Press, 2016

[4] R Hogervost and H Huijsing, *Design of Low-Voltage Low-Power Operational Amplifier Cells* Kluwer Academic Publishers, 1996

[5] R Hogervorst, "Compact Power-Efficient 3 V CMOS Rail-to-Rail Input/Output Operational Amplifier for VLSI Libraries," *IEEE J. Solid-State Circuits*, vol 29, no 12, pp 1505–1513, 1994

[6] L Langen and H Huijsing, "Compact Low-Voltage Power-Efficient Operational Amplifier Cells for VLSI," *IEEE J. Solid-State Circuits*, vol 33, no 10, pp 1482–1496, 1998

[7] G Ferri and W Sansen, "A Rail-to-Rail Constant-gm LowVoltage CMOS Operational Transconductance Amplifier," *IEEE J. Solid-State Circuits*, vol 32, no 10, pp 1536–1567, 1997

[8] J M Carrillo, J Duque-Carrillo, G Torelli, and J Ausin, "1-V Quasi Constant-gm Input/Output Rail-to-Rail CMOS Op-amp"," *Integr. VLSI J.*, vol 36, no 4, pp 161–174, 2003

[9] T Lehmann and M Cassia, "1-V power supply CMOS cascode amplifier," *IEEE J. Solid-State Circuits*, vol 36, no 7, pp 1082–1086, 2001

[10] E Raisanen-Ruotsalainen, K Lasanen, and J Kostamovaara, "A 1 2 V Micropower CMOS Op Amp with Floating-Gate Input Transistors," in *43rd IEEE Midwest Symp. Circuits and Systems*, 2000, pp 794–797

[11] B J Blalock, P E Allen, and G A Rincon-Mora, "Designing 1-V Op Amps Using Standard Digital CMOS Technology," *IEEE Trans. Circuits Syst. II*, vol 45, no 7, pp 769–780, 1998

[12] K Lasanen, E Raisanen-Ruotsalainen, and J Kostamovaara, "A 1-V 5 ,uW CMOS-Opamp with Bulk-Driven Input Transistors," in *43rd IEEE Midwest Symp. on Circuits and Systems*, 2000, pp 1038–1041

[13] H F Achigui, C J Fayomi, and M Sawan, "A DTMOS based 1 V OPAMP," in *Proc. ICECS*, 2003, pp 252–255

A Fast Alternating Direction Method of Multipliers Approach for 1-Bit Compressed Wideband SAR Imaging

M.DEMİR[1] and M. ARICI[2]

[1] Gaziantep University, Gaziantep/Turkey, mehmetdemir@gantep.edu.tr

Abstract **- Compressed Sensing (CS) theory emerged in recent years provide important benefits in Synthetic Aperture Radar (SAR) Imaging. Recent studies have shown that CS also provides good results in obtaining SAR images from 1-bit quantized measurements. In this paper, a new fast Alternating Direction Method of Multipliers (ADMM)-based algorithm is proposed for 1-bit CS SAR imaging. This proposed method is compared to ADMM-based algorithms previously used in this problem to demonstrate its effectiveness for 1-bit compressed SAR imaging**

Keywords **- SAR imaging from 1-bit data, Compressed Sensing, Alternating Direction Method of Multipliers (ADMM).**

I. INTRODUCTION

YNTHETIC aperture radar (SAR) is an imaging system that is able to capture high resolution image of earth in dark and adverse weather conditions. It is possible to increase image resolution by using wideband signals in range direction since there is a direct relation between transmitting signal bandwidth and image quality. However, wideband signal requires carrying a longer antenna which is impractical for aircraft and other moving systems and motivates one to use SAR imaging technology. On the other hand, the mentioned technology brings new burdens, for instance, more measurement requirement resulting with demand of faster analog-to-digital converters (ADC) and necessity of bigger storage space. The above given disadvantages leads us to study 1-bit quantization approach which addresses many benefits including low-cost SAR imaging opportunity. SAR imaging with 1-bit coded echo data is a subject previously studied in detail [1-3]. It is known that classical methods such as matched filtering (MF) used to obtain SAR images from 1-bit quantized data do not give very good imaging results.

The hardware that performs 1-bit quantization can be formed with a simple comparator. In this case, the speed of ADC is increased. On the other hand, since the magnitude information of the signal is lost in 1-bit quantization, it may be di cult to obtain a quality SAR image from 1-bit measurements.

The recently popular 1-bit compressed sensing (CS) framework was firstly presented in [4]. It is shown in [5] that stable signal recovery from 1-bit measurements is guaranteed. Since the important advantages of 1-bit quantization are known, many new methods for signal reconstruction using 1-bit data have been proposed. Binary iterative hard thresholding (BIHT) [5] and matching sign pursuit (MSP) [6] which are classified as greedy algorithms are suggested for solving 1-bit CS problems. Moreover, ℓ_1-norm minimization based renormalized fixed point iteration (RFPI) [4] and restricted-step shrinkage (RSS) [7] approaches that are not needed to know the sparsity level in advance are introduced. There are also convex method which is presented in [8] and Bayesian algorithm [9] for this optimization problem.

It is seen that 1-bit CS framework has also had applications for SAR imaging in recent years. A MAP-based method has been developed to obtain SAR image using 1-bit data [10]. In addition, a fast method that facilitates large vector-matrix products using Range-Doppler algorithm has been proposed for 1-bit SAR imaging [11].

In this study, a fast recovery algorithm based on Alternating Direction Method of Multipliers (ADMM), a specific type of Augmented Lagrangian Methods (ALM), is introduced to obtain a SAR image using 1-bit quantized measurements. ADMM-based approaches have been previously presented for 1-bit compressed SAR imaging [12, 13], but these produce slower solutions. The introduced method, which is computationally efficient, is important in terms of providing fast results for 1-bit compressed SAR imaging. In the presented work, the consistency between 1-bit data and recovered SAR signal is measured by sigmoid (logistic) function. Thus, an explicit unit-norm constraint is no longer required in optimization model [14]. Using logistic distribution as the term of data fidelity makes the problem convex and provides advantages in terms of optimization. While solving this problem with ADMM algorithm, minimization methods such as Newton or spectral conjugate gradient (SCG) were used to solve the unconstrained minimization caused by this term [13]. However, Newton method from these approaches complicates calculations since it needs Hessian operator in its numerical solution and SCG necessitates a lot of iterations to converge. In addition, oscillations that may occur while solving the cost function delay the minimization of problem. In order to overcome the aforementioned difficulties, a method that converges fast by suppressing oscillations and requires a simpler calculation

since it does not include Hessian matrix, is integrated into ADMM-based 1-bit CS SAR imaging algorithm [15].

II. One-bit CS Sar Imaging Formulation

A. SAR Observation Model

Since SAR observation model is based on Born approximation, SAR signal, re-turned from imaging scene, is obtained by multiplying target reflectivity by a linear operator. After target reflectivity field is converted to a lexicographic vector, SAR imaging model can be written in matrix form as follows:

$$\mathbf{r} = \mathbf{H}\mathbf{f} + \mathbf{n} \tag{1}$$

where $\mathbf{f} \in C^N$ is a vector representing target reflectivity, $\mathbf{r} \in \mathbb{C}^M$ is generated by demodulating SAR signal returned from field, $\mathbf{H} \in \mathbb{C}^{MxN}$ is a complex-valued mathematical model of SAR observation process, $\mathbf{n} \in \mathbb{C}^M$ is a complex-valued Gaussian noise vector.

The observed signal for CS SAR imaging framework is randomly sampled at sub-Nyquist rate in this work. Under-sampling scheme is formed by selecting randomly from fast-time and slow-time samples.

B. One-Bit Quantization

Measurements in in-phase and quadrature channels are quantized separately to 1-bit values in SAR imaging system.

$$\mathrm{w} = sign(\Re(\mathbf{r})) + j\sin(\Im(\mathbf{r})) \tag{2}$$

$\Re(.)$ and $\Im(.)$ show symbols of operators representing the real and imaginary parts of a vector, respectively. The $sign(.)$ function can be expressed as follows:

$$sign(x) = \begin{cases} +1 & x \geq 0 \\ -1 & x < 0 \end{cases} \tag{3}$$

In order to achieve objective function to be minimized, complex-valued vectors and matrices should be transformed into real-valued domain.

$$\mathbf{y} = \begin{bmatrix} \Re(\mathrm{w}) \\ \Im(\mathrm{w}) \end{bmatrix}, \ \mathbf{A} = \begin{bmatrix} \Re(\mathbf{H}) & -\Im(\mathbf{H}) \\ \Im(\mathbf{H}) & \Re(\mathbf{H}) \end{bmatrix}$$
$$\mathbf{x} = \begin{bmatrix} \Re(\mathbf{f}) \\ \Im(\mathbf{f}) \end{bmatrix}, \ \mathbf{e} = \begin{bmatrix} \Re(\mathbf{n}) \\ \Im(\mathbf{n}) \end{bmatrix} \tag{4}$$

After converting vectors and matrices into real-valued domain, equation (2) can be described as follows:

$$\mathbf{y} = sign(\mathbf{A}\mathbf{x} + \mathbf{e}) \tag{5}$$

In this equation, only the sign information of SAR measurements is retained. Our aim is to recover SAR image from this sign information. Consistency between **y** and **Ax+e** expressions is qualified by sigmoid function. The following metric [14] is obtained when consistency is enforced.

$$\phi(\mathbf{x}) = -\sum_{i=1}^{2M} \log(\sigma(y_i.(Ax)_i)) \tag{6}$$

where sigmoid function is $\sigma(x) = 1/(1+e^{-x})$. In order to reach a sparse solution, $\ell_1 - \text{norm}$ which is a convex relaxation of $\ell_2 - \text{norm}$ is usually utilized. Consistency and sparse penalty functions are combined to achieve the objective function where defined as following:

$$\begin{aligned} L(\mathbf{x}) &= -\sum_{i=1}^{2M} \log(\sigma(\mathbf{y}_i.(\mathbf{A}\mathbf{x})_i)) + \tau \|\mathbf{f}\|_1 \\ &= -\sum_{i=1}^{2M} \log(\sigma(\mathbf{y}_i.(\mathbf{A}\mathbf{x})_i)) + \tau \sum_{i=1}^{N} \sqrt{x_i^2 + x_{i+N}^2} \end{aligned} \tag{7}$$

The unconstrained optimization problem to be minimized is formulated as

$$\min_{x \in \mathbb{R}^{2N}} -\sum_{i=1}^{2M} \log(\sigma(\mathbf{y}_i.(\mathbf{A}\mathbf{x})_i)) + \tau \sum_{i=1}^{N} \sqrt{x_i^2 + x_{i+N}^2} \tag{8}$$

III. Fast ADMM Approach for 1-bit SAR Imaging

In order for solving the unconstrained optimization problem, which consists of two parts, in (8), a fast method based on ADMM has been suggested in this study. To make the unconstrained problem into constrained one, variable splitting is firstly applied:

$$\begin{aligned} &\min_{\mathbf{x} \in \mathbb{R}^{2N}, \mathbf{v} \in \mathbb{R}^{2N}} f_1(\mathbf{x}) + f_2(\mathbf{x}) \\ &\text{subject to } \mathbf{x} = \mathbf{v} \end{aligned} \tag{9}$$

where $f_1(\mathbf{x}) = -\sum_{i=1}^{2M} \log(\sigma(\mathbf{y}_i.(\mathbf{A}\mathbf{x})_i))$ and $f_2(\mathbf{v}) = \tau \sum_{i=1}^{N} \sqrt{x_i^2 + x_{i+N}^2}$.

The feature of ADMM approach is to try to solve the optimization problem separately by dividing it into small sub-problems. ADMM for this problem consists of the following sequential steps

$$\mathbf{x}_{k+1} = \arg\min_{\mathbf{x}} f_1(x) + (\mu/2)\|\mathbf{x} - \mathbf{v}_k - \mathbf{d}_k\|_2^2 \tag{10}$$

$$\mathbf{v}_{k+1} = \arg\min_{\mathbf{v}} f_2(\mathbf{v}) + (\mu/2)\|\mathbf{x}_{k+1} - \mathbf{v} - \mathbf{d}_k\|_2^2 \tag{11}$$

$$\mathbf{d}_{k+1} = \mathbf{d}_k - (\mathbf{x}_{k+1} - \mathbf{v}_{k+1}) \tag{12}$$

solving these minimization problems in sequential steps is easier than solving the initial unconstrained optimization problem

A. *New Approach for* x*-update Step*

For the purpose of obtaining a fast algorithm, a Newton-like optimization approach is utilized in x-update law. The method presented here can be seen as a second order expansion of steepest descent method both in time and space and it has various applications in mechanical (inelastic shocks) and control theory (asymptotic stability of oscillators, adaptive control)[16]. Let us formulate the optimization problem as a dissipative dynamical system with Hessian-driven damping. Namely,

$$\ddot{\mathbf{x}}(t)+\alpha\dot{\mathbf{x}}(t)+\beta\nabla^2 J(\mathbf{x}(t))\dot{\mathbf{x}}(t)+\nabla J(\mathbf{x}(t))=0 \tag{13}$$

where α and β are positive parameters. The model in (13) has some important features and is named as Dynamic Inertial Newton-like (DIN) method. Here, inclusion of momentum term $\alpha\dot{\mathbf{x}}(t)$ increases the rate of convergence dramatically and together with $\beta\nabla^2 J(\mathbf{x}(t))\dot{\mathbf{x}}(t)$ term, once we can find suitable values of and parameters so as to suppress transversal and longitudinal oscillations of solution on the valley axis of $J(\mathbf{x}(t))$ resulting with a faster algorithm. Besides, from the implementation point of view, it can be shown that DIN system in (13) behaves as a first order system (in time and space) where Hessian does not appear. Hence, complexity per iteration of the algorithm we aim to construct is further decreased. The aforementioned first-order system can be expressed as follows:

$$\begin{cases}\dot{\mathbf{x}}(t)+\beta\nabla J(\mathbf{x}(t))+\alpha\mathbf{x}(t)+b\mathbf{u}(t)=0\\ \dot{\mathbf{u}}(t)+\alpha\mathbf{x}(t)+b\mathbf{u}(t)=0\end{cases} \tag{14}$$

where a, b and β are positive constants. The system given in (14) is called generalized DIN (GDIN). If the cost function J is convex and $a=\alpha-1/\beta, b=1/\beta$ conditions are met, DIN and GDIN systems are equivalent in numerical application[15]. Based on these features, the numerical fast update algorithm for x can be written as follows:

$$\begin{cases}\mathbf{x}_{k+1}=\mathbf{x}_k-\beta\nabla J(x_k)-a\mathbf{x}_k-b\mathbf{u}_k\\ \mathbf{u}_{k+1}=\mathbf{u}_k-a\mathbf{x}_k-b\mathbf{u}_k\end{cases} \tag{15}$$

Remark. When choosing β parameter here, the sampling interval should be considered and not to be selected too high. Taking broad steps may cause the solution to have difficulty in catching the global minimum. On the other hand, the most important criterion in choosing a and b parameters is that they are greater than 0. While selecting these parameters in the algorithm, Monte Carlo simulation was used to provide the best performance.

B. *Computing* v*-update Step*

The optimization problem (11) can be regarded as Moreau envelope of f_2. It can be formulated as following:

$$\psi_{\tau\phi/\mu}(\mathbf{z})=\arg\min_{\mathbf{v}}\ \frac{\tau}{\mu}\sum_{i=1}^{N}\sqrt{\mathbf{v}_i^2+\mathbf{v}_{i+N}^2}+\frac{1}{2}\|\mathbf{v}-\mathbf{z}\|_2^2 \tag{16}$$

where $\mathbf{z}=\mathbf{x}_{k+1}-\mathbf{d}_k$. If the real and complex parts of $\mathbf{v}$ are considered as two separate variables for i = 1, 2…, N, N minimization problems can be written as

$$\begin{pmatrix}\psi_{\tau\phi/\mu}(\mathbf{z}_i)\\ \psi_{\tau\phi/\mu}(\mathbf{z}_{i+N})\end{pmatrix}=\arg\min_{\varsigma,\xi}\ \frac{\tau}{\mu}\sum_{i=1}^{N}\sqrt{\varsigma_i^2+\xi^2}+\frac{1}{2}(\varsigma-\mathbf{z}_i)^2+\frac{1}{2}(\xi-\mathbf{z}_{i+N})^2 \tag{17}$$

The following equation expresses the closed form solution of (17) for i = 1, 2, …..,N.

$$\begin{aligned}\psi_{\tau\phi/\mu}(\mathbf{z}_i)&=(\sqrt{\mathbf{z}_i^2+\mathbf{z}_{i+N}^2}-\frac{\tau}{\mu})_+\frac{\mathbf{z}_i}{\sqrt{\mathbf{z}_i^2+\mathbf{z}_{i+N}^2}}\\ \psi_{\tau\phi/\mu}(\mathbf{z}_{i+N})&=(\sqrt{\mathbf{z}_i^2+\mathbf{z}_{i+N}^2}-\frac{\tau}{\mu})_+\frac{\mathbf{z}_{i+N}}{\sqrt{\mathbf{z}_i^2+\mathbf{z}_{i+N}^2}}\end{aligned} \tag{18}$$

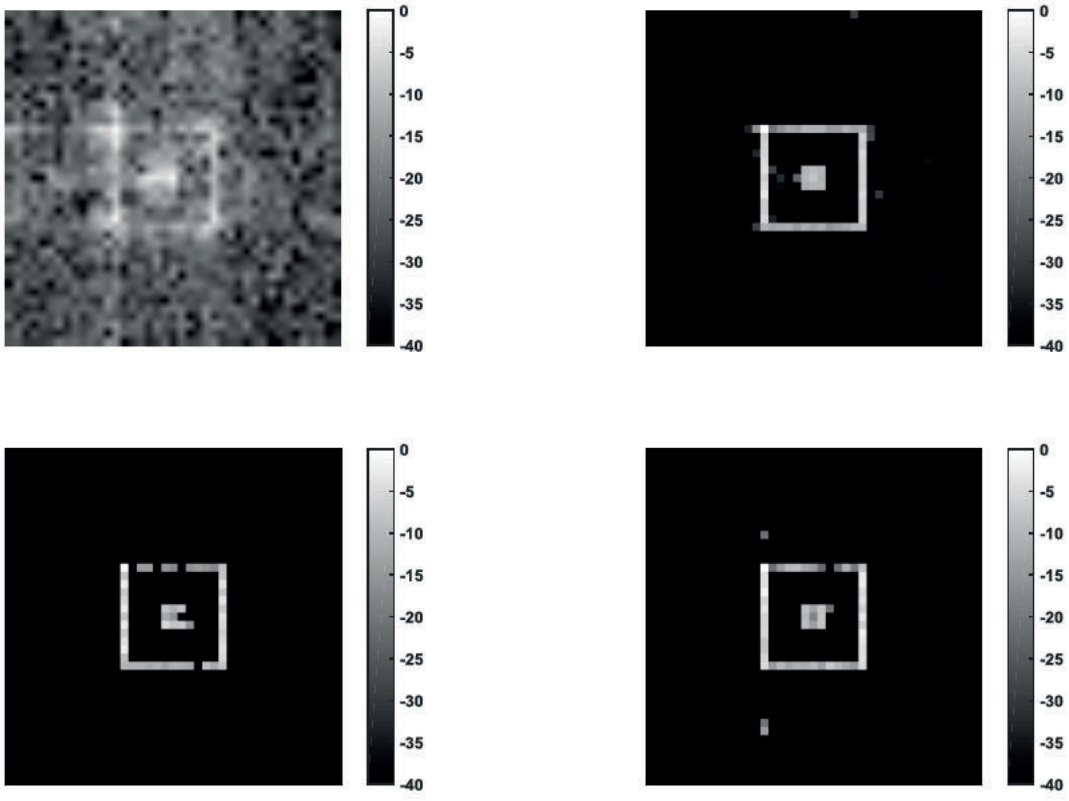
$(\gamma)_+=\max(\gamma,0)$. With this solution, by combining the x-update step that is solved using GDIN method, the proposed ADMM-GDIN emerges as shown in the following **Algorithm**.

Algorithm: ADMM-GDIN Algorithm for 1-Bit CS SAR

1. Choose $\mu>0$, $\tau>0$. Initialize **v**, **d**
2. k=0
3. **Repeat:**
4. $\mathbf{x}_{k+1}$ = GDIN(**A**, **y**, **d**, **v**, μ). GDIN function is called.
5. $\mathbf{d}_{k+1}=\psi_{\tau\phi/\mu}-(\mathbf{x}_{k+1}-\mathbf{d}_k)$
6. $\mathbf{d}_{k+1}=\mathbf{d}_k-\mathbf{x}_{k+1}+\mathbf{v}_{k+1}$
7. $k\leftarrow k+1$

IV. Simulation Results

In this section, two simulations have been carried out to demonstrate whether the proposed approach is effective or not in 1-bit SAR imaging. In the first of these two simulations, synthetically produced sparse SAR scene is used. In the second simulation, compared reconstruction algorithms have been performed on a SAR data extracted from a real SIR-C/X-SAR image. In the experiments, ADMM-NM, ADMM-SCGM and proposed ADMM-GDIN approaches have been tested. In addition, SAR image obtained from 1-Bit MF reconstruction method has been added to simulations in order to see clearly that high side-lobes are eliminated in the introduced approach. Regularization parameter has been set manually and the same value is entered for all methods. Simulations have been performed at SN R = 20 dB level. SAR system parameters can be seen at Table 1.



© (d)

Fig. 1 Log-intensity images obtained from four different methods using synthetically produced SAR scene under SNR=20 dB: (a) 1-Bit MF, (b) ADMM-NM, (c) ADMM-SCGM, (d) The proposed ADMM-GDIN approach.

Table 1: SAR System Parameters

Parameters	Values
[illegible]	[illegible]
[illegible]	[illegible]
[illegible]	

When obtained SAR images are compared, it has been seen that all ADMM-based methods suppress side-lobes and remove ghost targets, which is an important problem for 1-bit SAR. It is clearly seen from Figure 1 that the proposed ADMM-GDIN algorithm produces similar results with other two ADMM-based algorithms.

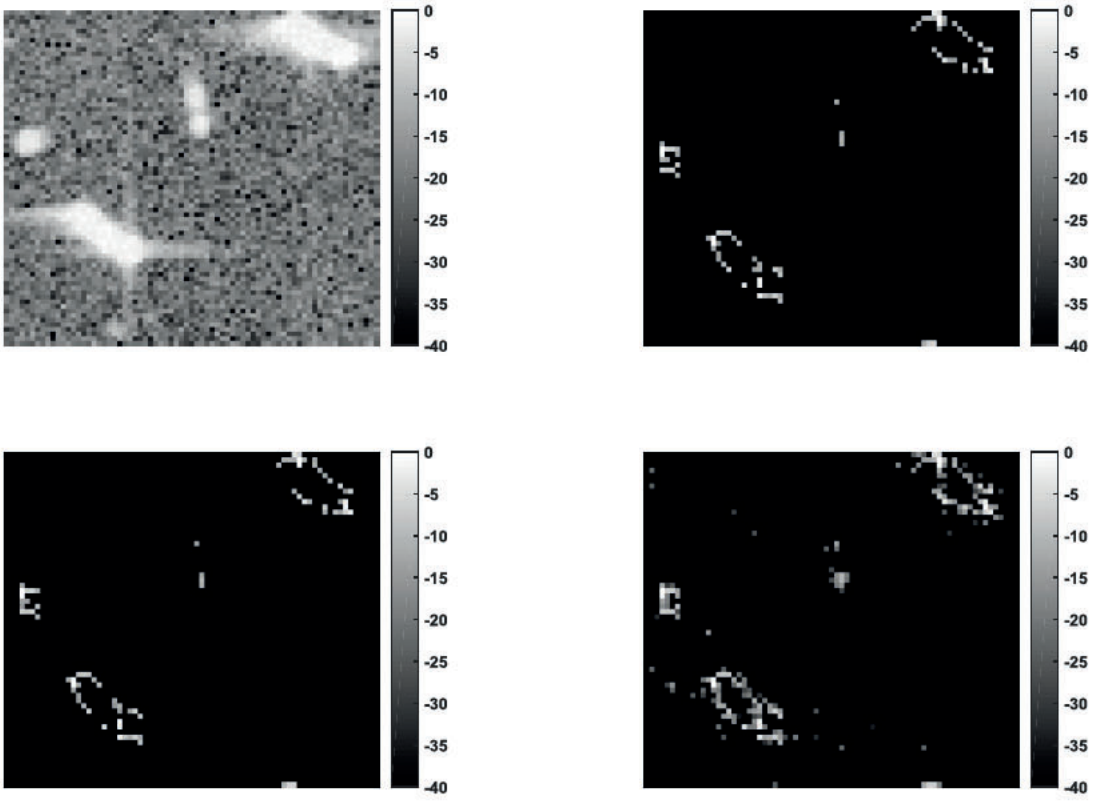
In the second experiment, phase history data has been produced using some part of a real SIR-C/X-SAR image. In this experiment, the results acquired by four methods are compared as shown in Figure 2. It is seen that high side-lobes are removed in SAR images obtained from ADMM-based approaches. However, in SAR image produced by 1-Bit MF algorithm, high side-lobes attract attention. Figure 2 demonstrates that target-to-background contrast increased too much in SAR images generated by ADMM-based approaches. = 130 and = 1 are selected in this simulation. In ADMM-NM, ADMM-SCGM and ADMM-GDIN methods, SAR images are produced using 22:8% of all data.

Lastly, the run-time comparisons of ADMM-NM, ADMM-SCGM and ADMM-GDIN for both experiments are given in Table 2. Simulations are performed on a workstation with Intel Xeon 3.50 GHz processor and 24 GB of RAM. As seen from Table 2, ADMM-GDIN, which is a computationally efficient algorithm, produces faster results than ADMM-NM and ADMM-SCGM. The run-time difference between ADMM-GDIN and other two algorithms appears to be too large.

Table 2: Run-time Comparisons

V. Conclusion

In this study, a new approach named as ADMM-GDIN is presented for 1-bit compressed SAR imaging. In the x-update step of ADMM-GDIN algorithm, the problem is minimized



(c) (d)

Fig. 2 Log-intensity images obtained from four different methods using SAR data extracted from SIR-C/X-SAR image under SNR=20 dB: (a) 1-Bit MF, (b) ADMM-NM, (c) ADMM-SCGM, (d) The proposed ADMM-GDIN approach.

using fast GDIN method. By integrating this algorithm into ADMM, a new approach that produces very fast results for 1-bit compressed SAR imaging has been developed. The proposed approach has been shown to be effective in simulations with data generated using synthetic and real SAR scenes. As can be seen from comparisons, the proposed algorithm yields very fast results without impairing the SAR image quality. The proposed ADMM-GDIN can be used as an alternative previously available methods for 1-bit CS SAR.

REFERENCES

[1] G Franceschetti, V Pascazio, and G Schirinzi, "Processing of signum coded sar signal: theory and experiments," *IEE Proceedings F - Radar and Signal Processing*, vol 138, no 3, pp 192–198, 1991

[2] G Franceschetti, S Merolla, and M Tesauro, "Phase quantized sar signal processing: theory and experiments," *IEEE Transactions on Aerospace and Electronic Systems*, vol 35, no 1, pp 201–214, 1999

[3] G Alberti, G Franceschetti, G Schirinzi, and V Pascazio, "Time-domain convolution of one-bit coded radar signals," *IEE Proceedings F - Radar and Signal Processing*, vol 138, no 5, pp 438–444, 1991

[4] P T Boufounos and R G Baraniuk, "1-bit compressive sensing," *in 2008 42nd Annual Conference on Information Sciences and Systems*, Princeton, NJ, USA, 2008, pp 16–21

[5] L Jacques, J N Laska, P T Boufounos, and R G Baraniuk, "Robust 1- bit compressive sensing via binary stable embeddings of sparse vectors," *IEEE Transactions on Information Theory*, vol 59, no 4, pp 2082–2102, 2013

[6] P T Boufounos, "Greedy sparse signal reconstruction from sign measurements," *in 2009 Conference Record of the Forty-Third Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, USA, 2009, pp 1305–1309

[7] J N Laska, Z Wen, W Yin, and R G Baraniuk, "Trust, but verify: Fast and accurate signal recovery from 1-bit compressive measurements," *IEEE Transactions on Signal Processing*, vol 59, no 11, pp 5289–5301, 2011

[8] J N Laska, Z Wen, W Yin, and R G Baraniuk, "Trust, but verify: Fast and accurate signal recovery from 1-bit compressive measurements," *IEEE Transactions on Signal Processing*, vol 59, no 11, pp 5289–5301, 2011

[9] J N Laska, Z Wen, W Yin, and R G Baraniuk, "Trust, but verify: Fast and accurate signal recovery from 1-bit compressive measurements," *IEEE Transactions on Signal Processing*, vol 59, no 11, pp 5289–5301, 2011

[10] X Dong and Y Zhang, "A map approach for 1-bit compressive sensing in synthetic aperture radar imaging," *IEEE Geoscience and Remote Sensing Letters*, vol 12, no 6, pp 1237–1241, 2015

[11] C Zhou, F Liu, B Li, J Hu, and Y Lv, "A 1-bit compressive sensing approach for sar imaging based on approximated observation," *in Eighth International Conference on Digital Image Processing (ICDIP 2016)*, vol 10033, 2016, pp 100 333J–1–100 333J–7

[12] M Demir and E Erçelebi, "One-bit compressive sensing with time-varying thresholds in synthetic aperture radar imaging," *IET Radar, Sonar & Navigation*, vol 12, no 12, pp 1517–1526, 2018

[13] M Demir and E Erçelebi, "Admm with spectral conjugate gradient method for 1-bit compressed sar imaging," *in 2019 9th International*

Conference on Recent Advances in Space Technologies (RAST), Istanbul, Turkey, 2019, pp 423–428

[14] J Fang, Y Shen, H Li, and Z Ren, "Sparse signal recovery from one-bit quantized data: An iterative reweighted algorithm," *Signal Processing*, vol 102, pp 201 – 206, 2014

[15] F Alvarez, H Attouch, J Bolte, and P Redont, "A second-order gradient-like dissipative dynamical system with hessian-driven damping : Application to optimization and mechanics," *Journal de Mathématiques Pures et Appliquées*, vol 81, no 8, pp 747 – 779, 2002

[16] M Arıcı and T Kara, "Improved adaptive fault-tolerant control for a quadruple-tank process with actuator faults," *Industrial & Engineering Chemistry Research*, vol 57, no 29, pp 9537–9553, 2018

ADJUSTMENT OF PHASE ARRAY RADARS USİNG TYPE-2 FUZZY LOGİC

Fatih Turgut AKÇAY[1] and İsmail SARITAŞ[2]

[1] Selcuk University, Konya/Turkey, fatihturgutakcay@gmail.com
[2] Selcuk University, Konya/Turkey, isaritas@selcuk.edu.tr

Abstract - Today, radar technology is becoming increasingly important in both civilian and military fields. With the development of the defense industry, studies on radars are increasing and the importance of radar systems becomes more prominent. The national radar development project for our country is also one of the important studies emphasizing this importance for air defense systems and radar technologies.

In this thesis study, the enemy aircraft entering the airspace under the conditions of the field are detected and identified with the help of type-2 fuzzy logic and the target's distance, speed, altitude, acceleration and radar cross section area are used.

It is aimed to prioritize and optimize the time sharing of the radar, which is necessary for the air threats for the efficient operation of the radars.

In order to make this optimization, the determined parameters will be processed using type-2 fuzzy logic and the output value will be taken.

In this study, the advantages of type-2 fuzzy logic over type-1 fuzzy logic and its success in time management are examined.

Thanks to this thesis study, it is thought that the efficiency of the radars produced / planned with the resources of our country will be increased in the defense industry in the coming years.

Keywords - Fuzzy Logic, Phased Array Radar, Radar Technologies, Defense Industry, Classification, Type-2 Fuzzy Logic

I INTRODUCTION

With the start of the Second World War, during the "British Wars" between Germany and England, the British actively used radar technology to protect their country against German air strikes. Along with the developing technology and methods used in time, the usage areas and capabilities of radars have also improved. Today, the radars of the countries are used for air surveillance, search, detection and fire control radars in ADS (Air Defense System) for the security of their airspace.

In today's war conditions, radars have an indispensable importance for airborne attacks and it is very important for the operating environment to follow each enemy aircraft in the airspace and to launch missiles within the firing range when necessary.

Radar resources should be used efficiently in order to track the many and various threats found in the airspace.

Some parameters of air threats should be examined for the most effective use of the radar source. In the study, five kinds of parameters were taken into consideration in determining the criticality level of the air platform for the radar. These parameters are the speed of the threat, the altitude of the threat, the distance of the threat, the acceleration of the threat and the radar cross section (RCS) of the threat. By looking at these parameters, the TTT (Target Tracking Time) value, which determines how long the radar will follow the air threat, will be revealed with the help of type-2 fuzzy logic.

When the literature studies are examined, it is seen that there are many studies using fuzzy logic and involving the defense industry.

In the study using fuzzy logic, it was tried to measure the threat criticality level by evaluating the characteristics such as the target type, speed, angle of arrival, flight altitude, and whether it has the ability to jam [1].

In another study conducted with fuzzy logic, the threat was graded on the score scale prepared using the speed, altitude, distance information and the threat was classified according to the criticality level. In the study, it was not examined whether there was any electronic support system on threats that only speed, altitude and distance information were evaluated [2].

In the study using type-1 fuzzy logic, warships should perceive the threats coming from the air in a limited time and show the appropriate reaction because the aircraft move faster than land or sea vehicles, radar, electronic support systems, eye It evaluates that information from various sensors such as, is used, determining the target type and threat level is of great importance in terms of the reaction to be shown [3]. In this study, the threat level was determined using fuzzy logic rules, which were formed by taking expert opinion with the values of speed, altitude, distance and climb rate of air targets. They applied fuzzy logic simulation results to triangular and gaussian distributions using Sugeno and Mamdani techniques and compared the results [4,5].

In another fuzzy logic study, it optimized the radar source and followed the targets with higher priority in the operational field for a longer period of time, thus enabling more threats with high danger to be followed and kept under surveillance [6].

In studies conducted with fuzzy logic, the speed and angle information of the target was used with fuzzy logic to calculate how far the missile would explode to the target and the best explosion distance was found [7].

When the threat started to maneuver by using the radar fuzzy logic, it increased the time to look at that threat and followed it for a longer time and provided efficiency in the radar source [8].

By examining the parameters of the radar, coefficients were given to the parameters with certain changes, and a study was conducted to classify these parameters with the help of fuzzy logic and to use the radar source effectively [9].

Resource allocation optimization of phased array radars was discussed and Lagrange relaxation model was used [10,11]. According to the Lagrange relief model, the search and tracking features of the phased array radar are considered independently from each other. Resource constraints converted to integer linear problems and optimized.

By making the definitions of type-2 fuzzy logic clusters, he talked about their advantages over type-1 fuzzy logic clusters and a modeling study of Fuzzy Tool Box on SIMULINK was performed in MATLAB software environment [12].

By investigating the difficulty of making decisions in a fuzzy environment, a study has been carried out to find that type-2 fuzzy logic will give more accurate results than type-1 fuzzy logic, and a suitable and correct supplier selection study has been carried out using type-2 fuzzy logic [13].

II PHASE ARRAY RADARS

In conventional radar, the direction of the antenna is changed to try to increase the area seen by the radar. Thus, the area seen by the radar and the surveillance efficiency are increased. Since the power supply of the active elements that make up the array is separate in phase array radars, the beam amplitude and frequency are directed electrically. Without rotating the antenna mechanically, the phase difference is added between the frequency transmissions from the transmitter antenna and the antenna is rotated electrically. The advantage of a phased array over a conventional radar is that the frequency beam can be steered without any mechanical system. In this way, mechanical errors are eliminated and radar systems become more reliable. Higher scan rates are obtained as beam positioning becomes faster.

Antenna outputs have separate power feeds so it can have multiple beams at the same time and can track multiple threats simultaneously.

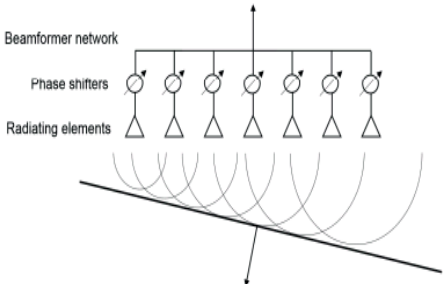


Figure 1: Phase array antenna structure with phase difference applied.

Another advantage of phased array radars is that the beam size transmitted can be changed.

If parameters such as maneuver of threats and speed are not fully detected while scanning the airspace, it can change the length or width of the phased array radar beam and update itself to catch the target. Antennas that work with separate feeding power placed on the array antenna can be raised to high power levels, so threats at longer distances can be followed. Another feature of phase array radars is that they solve the clutter situation, which is expressed as clutter or chaos, more effectively than conventional radars. Electronic scanning capability provides great advantages and competencies to phased array radars.

Considering all these features, phased array radars have been used in history since the 1950s. Considering the first use of phased array radars, radars were used as more than one broadcast source to better view the picture of the airspace in the late 1940s. In the 1950s, using electronic card designs, multiple broadcast sources were collected on a single antenna. This new design enables faster scanning to more areas in the sky electrically without mechanically rotating the antenna.

III FUZZY LOGIC

Fuzzy logic emerged with a study conducted in 1961 by the Azerbaijani scientist Lotfi A. Zadeh. Zadeh published a publication aiming to solve systems that are difficult to model in mathematics by putting them in logic. Fuzzy logic method was used to make decisions easily in situations that are difficult to explain with sharp expressions in mathematical models.In the early years of the 1970s, the use of fuzzy logic was used in many areas such as economy, medicine, decision-making mechanisms, factory automation.

According to the set theory, which was founded by Aristotle in the 4th century BC and called classical logic, each of the objects in space must be a member of the set defined or not. In logic clusters, this situation is called Logic 1 if the cluster element is called Logic 0 if the cluster element is not. In this way, the clusters whose belonging states are determined as 0 or 1 are called crisp clusters.

In fuzzy logic clustering, the elements belonging to the cluster do not belong to the cluster in the form of 0 or 1 like classical clusters, but they belong to a set with a value between 0 and 1. In other words, an element is a set member of a set to a certain extent.

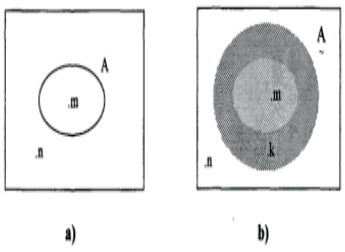


Figure 2: (a) Classical set (b) Fuzzy set [14]

With fuzzy logic, verbal variables were used in decision mechanisms. We can use verbal values to express the age of individuals, for example, the individual is young, very young, not very young, not young, old, not too old, middle-aged. Using fuzzy logic, we can define these verbal expressions as elements of fuzzy sets. In this way, we can create decision mechanisms by using verbal expressions determined by non-numeric people as input data.

Type-1 fuzzy logic systems have been used since the 1970s to bring uncertain systems to certain results. It has been observed that sometimes Type-1 fuzzy logic systems are insufficient in system models used to make the concept of uncertainty specific, and Type-2 fuzzy logic systems have been developed and started to be used in order to reach more precise results.

Unlike traditionally in use Type-1 fuzzy logic clusters, Type-2 fuzzy logic clusters form an extended cluster with lower and upper layers at their boundaries. In Type-2 fuzzy logic sets, the boundaries of the memberships of the cluster elements are not precise, unlike Type-1. In other words, membership transfers are made within certain limits. Since the membership function is transitive within certain limits, uncertainties in the rules of the fuzzy logic system are avoided.

There are some disadvantages in Type-1 fuzzy logic systems because the limits of membership functions are sharp and specific. To give an example of these disadvantages, the words used in the creation of the rule have different meanings between different people, the differences between the expert opinions consulted when determining the membership functions, the difficulty in establishing the membership function boundaries, and the noise rates in the data used to establish the fuzzy system with the input data can be given examples of such disadvantages.

The sharpness at the boundaries of type-1 fuzzy logic systems cannot fully model the cluster when membership transitions are not clear and may produce erroneous results. Since the boundaries of membership functions in Type-2 fuzzy logic systems are transitive, even in cases of uncertainty and failure, they can give more accurate results than Type-1 fuzzy sets.

When fuzzy logic systems are considered, it is seen that the rules affecting the outputs of the system consist of parts such as fuzzy stage, inference stage and clarification and give an output according to rule bases. In order for the fuzzy logic system to give an output, it must work with a function such as $y = f(x)$. Fuzzy logic systems that produce outputs according to function logic are used in many sectors today, such as control systems, production processes, detection systems, etc. There are field applications in the form of.

In fuzzy logic systems, IF - THEN rules are required to reach an output result from input data. When the rule structure is examined, the section expressed as IF can be expressed as the section where the input data is assigned, the section expressed as ISE can be expressed as the section where the output result is assigned. This set of rules can be used for nth order fuzzy logic system without any type limitation. High-order fuzzy logic systems are not preferred much because the number of parameters to be used will increase as the degree increases. In the Type-1 fuzzy logic system, membership functions are expressed with absolute values, whereas in the Type-2 fuzzy logic system, the membership function itself contains fuzzy boundaries.

If naming is to be made in fuzzy logic systems, the degree to which the system includes mostly fuzzy logic system is checked. For example, if the fuzzy logic system is defined with Type-1 fuzzy logic systems, it is called a Type-1 fuzzy logic system. If a fuzzy logic system has both Type-1 and Type-2 fuzzy clusters, this system is called Type-2 fuzzy logic system.

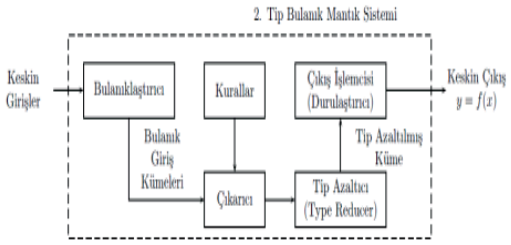


Figure 3: Type-2 fuzzy logic block diagram

IV RADAR MODELING WITH TYPE-2 FUZZY LOGIC

A. Problem identification and solution technique

In this thesis, the working logic of phased array radars and Type-2 fuzzy logic systems are mentioned. It is observed that in today's battlefields,

weapon systems and ammunition have become much more electronic, complex and intelligent, and weapon systems and parts used more than body strength in the battlefield have become affecting the outcome of the war. When examined in this respect, for the independence of countries, they have to protect their airspace and support this independence with advanced systems.

In this study, using the limited time resource in the most efficient way and ensuring the security of the airspace by reducing the target tracking times (TTT) of the Type-2 fuzzy logic systems and phased array radars compared to the Type-1 fuzzy logic systems, the weighted coefficient systems and the fixed coefficient systems. intended.

For the work done, http://ritweb.cloudapp net:8080/JuzzyOnline web address was used and system design, rule definitions and outputs were calculated using the online web page.

In the operation of the radar, there is a direct proportion between the change in the amount of TTT and the radar's ability to detect the air target. For example, when a radar with constant coefficient is following the target, when the air target accelerates or accelerates by maneuvering to the right and left, the radar may have difficulty in tracking the target with the changes that may occur in the target. If these parameters change too quickly, it may become impossible for the radar to catch or detect the target after a while. For this reason, the more targets are tracked, the more parameter changes such as the position, speed and acceleration of the targets are observed and the rate of target detection is increased. In accordance with the purpose of the thesis study, since the advantages of the Type-2 fuzzy logic system compared to the Type-1 fuzzy logic system, the systems with weighted coefficients and the systems with constant coefficients were examined, the target detection ratios were taken equal for all systems. Even if the statistics for detecting targets of the radars are required as 1 in practice, this rate is not possible in practice, and today the most ideal target detection rates vary between 0.85 and 0.95.

B. Design of the radar using a Type-2 fuzzy logic system

While designing the Type-2 fuzzy logic system, five types of data used by phased array radars to detect the target were evaluated. These data are the distance of the threat, the speed of the threat, the altitude of the threat, the acceleration of the threat and the radar cross section of the threat. While designing the radar, the range that the radar can detect is 400 km, the speed it can follow is 1000 m / sec, the altitude it can detect is 60,000 feet, the highest acceleration it can follow is 9 G, so that it can be compared with the Type-1 fuzzy logic system and other systems that were previously studied. the radar cross section (RCS - Radar Cross Section) it can take has been selected as 10 m2.

The distance of the threat from the radar is inversely proportional to the priority of the threat, and a distant threat is less important to our radar. As the distance between the threat and the radar decreases, the significance of the threat increases and the amount of TTT increases.

For the phased array radar modeled using Type-2 fuzzy logic, membership functions are determined for each parameter, that is, data entry. Here, linguistic expressions are used when determining membership functions. While choosing linguistic expressions, attention has been paid to the use of words that do not create meaning confusion and can have the same effect on everyone when they are said.

The linguistic expressions of the designed Type-2 fuzzy logic phased array radar are as in Table 1.

Table 1: Linguistic expressions table

Parameter	Linguistic Expression		
Altitude	Low	Middle	High
Speed	Small	Middle	Fast
Range	Near	Middle	Far
Acceleration	Low	Normal	High
RCS	Small	Middle	Big

The system is generally designed as a system that tries to bring the amount of radar TTT, which is 243 rules, whose entrance membership functions are gaussian and the exit membership function is triangular, with 5 inputs and 1 output, to the most efficient level. General system view showing system inputs and outputs is shown in Figure.

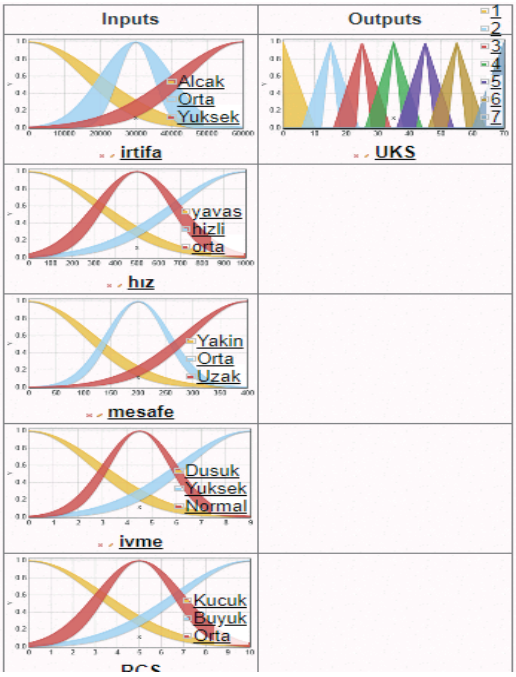


Figure 4: Type-2 fuzzy logic system general view

When establishing a type-2 fuzzy logic system, since each of the 5 input data selected as input parameters has 3 membership functions, the system has been defined with 3 ^ 5 = 243 IF-THENs. The weights of the parameters selected for the input are not different from another parameter, which has been tried to be kept as close to each other as possible in order to ensure consistency between the results.

There are Type-1 fuzzy logic systems in Type-2 fuzzy logic as mentioned before. In order to get an output result in Type-2 fuzzy logic systems, it is necessary to reduce the system to the Type-1 level first. In the design of our radar, the center of mass type reduction method was used as the type reduction method.

For the output result of the fuzzy logic system reduced to Type-1 level, the centroid (center of gravity) clarification technique was chosen.

V TEST RESULTS OF USING THE RADAR WITH TYPE-2 FUZZY LOGIC

For the test of the radar system created using the Type-2 fuzzy logic system, the success of the system in resource management was measured by examining the scenarios of increasing and decreasing the amount of TTT, and it was examined whether the resource management was used efficiently compared to the similarly established Type-1 systems.

The table below examines the change in the amount of TTT of the radar as the altitude level changes.

Table 2: Scenario where altitude is variable

Altitude	Speed	Range	Acceleration	RCS	Result (sec)
55000	999	399	8.9	0.1	36.06
45000	999	399	8.9	0.1	40.43
35000	999	399	8.9	0.1	46.74
25000	999	399	8.9	0.1	49.52
15000	999	399	8.9	0.1	52.39
5000	999	399	8.9	0.1	54.29

As can be seen from the table, when the altitude value of the target decreases until the speed, distance, acceleration and RCS values are kept constant, the TTT value of the radar operating with the type-2 fuzzy logic system increases. For a target with a high altitude, the radar can follow the target for less time, using time resources for other threats and increasing efficiency.

In the table below, the change in the TTT amount of the radar as the distance level changes is examined.

Table 3: Scenario where range is variable

Altitude	Speed	Range	Acceleration	RCS	Result (sec)
55000	999	399	8.9	0.1	36.06
55000	999	350	8.9	0.1	36.18
55000	999	300	8.9	0.1	41.23
55000	999	250	8.9	0.1	46.49
55000	999	200	8.9	0.1	49.35
55000	999	150	8.9	0.1	51.38
55000	999	100	8.9	0.1	53.62
55000	999	50	8.9	0.1	54.66
55000	999	10	8.9	0.1	54.74

As can be seen from the table, when the distance value of the target whose altitude, speed, acceleration and RCS values are kept constant, the TTT value increases continuously. As it can be understood from here, by tracking the target at a long distance for less time, it can use the time resource for other threats and increase efficiency.

In the table below, the change in the TTT amount of the radar as the speed level changes is examined.

Table 4: Scenario where speed is variable

Altitude	Speed	Range	Acceleration	RCS	Result (sec)
55000	100	399	8.9	0.1	26.13
55000	200	399	8.9	0.1	28.09
55000	300	399	8.9	0.1	29.93
55000	400	399	8.9	0.1	31.26
55000	500	399	8.9	0.1	32.20
55000	600	399	8.9	0.1	32.97
55000	700	399	8.9	0.1	33.96
55000	800	399	8.9	0.1	35.08
55000	900	399	8.9	0.1	35.72
55000	999	399	8.9	0.1	36.06

As can be seen from the table, it is observed that there is an increase in the amount of TTT when the target speed increases continuously by keeping the altitude, distance, acceleration and RCS values constant. As it can be understood from here, by staying above a target with increasing speed, we will be able to follow the target and use the air defense system more effectively.

The table below examines the change in the TTT amount of the radar as the acceleration level changes.

Table 5: Scenario where acceleration is variable

Altitude	Speed	Range	Acceleration	RCS	Result (sec)
55000	999	399	1	0.1	25.31
55000	999	399	2	0.1	27.72
55000	999	399	3	0.1	30.32
55000	999	399	4	0.1	31.93
55000	999	399	5	0.1	32.91
55000	999	399	6	0.1	33.81
55000	999	399	7	0.1	35.17
55000	999	399	8	0.1	35.84
55000	999	399	8.9	0.1	36.03

As can be seen from the table, it is observed that there is an increase in the amount of TTT of a target whose acceleration is constantly increased by keeping the altitude, speed, distance and RCS values constant. Considering that a target with increased acceleration poses more danger for radar, it is considered that more target tracking is required.

In the table below, the change in the amount of TTT of the radar as the RCS value changes is examined.

Table 6: Scenario where RCS is variable

Altitude	Speed	Range	Acceleration	RCS	Result (sec)
55000	999	399	8.9	9.9	25.08
55000	999	399	8.9	9	26.20
55000	999	399	8.9	8	28.15
55000	999	399	8.9	7	29.98
55000	999	399	8.9	6	31.29
55000	999	399	8.9	5	32.22
55000	999	399	8.9	4	32.99
55000	999	399	8.9	3	33.98
55000	999	399	8.9	2	35.09
55000	999	399	8.9	1	35.73
55000	999	399	8.9	0.1	36.06

As can be seen from the table, it is observed that there is an increase in the amount of HTS of a target whose RCS value is constantly decreased by keeping the altitude, speed, distance and acceleration values constant. Based on this, it is considered that it would be beneficial to follow the target for a long time since the decrease in the surface area of jet aircraft, which have smaller RCS values compared to commercial passenger aircraft, will make it difficult to track the target.

In the table below, the comparison of the radar scenarios designed according to the Type-1 fuzzy logic system previously prepared by Er with the Approximate Type-2 fuzzy logic systems is examined [15]. Since Er did not use the RCS value

in the radar design he prepared with type-1 fuzzy logic, this value was accepted as the average value of 5 in order to minimize the coefficient and effect [15].

Table 7: Differences between type-1 and type-2 systems.

Range	Altitude	Speed	Acceleration	RCS	Type-1	Type-2
211 880	58441	529 7	4 71	5	55 8	28 61
166 994	52231	417 485	3 71	5	59 5	33 06
210 771	53910	526 928	4 68	5	55 0	28 49

As can be seen from the table, in radar designs with two fuzzy logic scenarios designed under similar conditions, the Type-2 fuzzy logic system produces values under the Type-1 fuzzy logic system in terms of time and allows the remaining time to be used for scanning the airspace or tracking other threats.

VI RESULT

One of the greatest guarantees of the independence of countries is the security of airspaces. Air defense systems, which are used to protect the security of airspaces, are of undeniable importance. In this study, it is aimed to use the limited time resource in the best way by optimizing the tracking times of the targets by using Type-2 fuzzy logic in order to use air defense systems in the most effective way, and to reach shorter follow-up times compared to the similarly designed Type-1 fuzzy logic. In the tests, it was concluded that the phased array radar system operating with Type-2 fuzzy logic produced shorter results than the Type-1 and followed the threats in the ideal time periods according to the evaluated parameters of the target.

It is considered that it would be appropriate to test fuzzy logic modeling in real field environment in order to give more accurate and sharper results in test simulations. In addition to the parameters examined in this study, system sensitivity can be increased by increasing the number of parameters to be considered or by integrating intelligence information from other systems.

REFERENCES

[1] Lu Y , Wang Y , Lei Y , " Air target theat assesment based on fuzzy rough reasoning", The 27th Chinese Control and Decision Conference (2015 CCDC), Qingdao, 885-890, 2015

[2] Rizwan H , Tayyaba S , Ashraf M W , Rasheed H , Ahmed Z , Ali B , "Threat evaluation of suspicious target for cognitive radar", 17th IEEE International Multi Topic Conference 2014, Karachi, 456-461,2014

[3] Beser, F , Adıgüzel, D , Yıldırım, Ö , Yıldırım, T , Bulanık Mantık Kullanarak Hava Savunma Karar Destek Sistemi Tasarımı, Akıllı Sistemler ve Uygulamaları Dergisi, Cilt: 1, Sayı: 1, Sayfa 135-139, 2018

[4] Sugeno M , "Industrial applications of fuzzy control", Elsevier Science Pub Co , 1985

[5] Mamdani E H , Assilian S , "An expriment in linguistic synthesis with a fuzzy logic controller", International Journal of Man-Machine Studies, 7(1), 1-13, 1975

[6] Stoffel, A P , Heuristic Energy Management for Active Array Multıfunction Radars, IEEE National Telesytem Conference, 71p , 1994

[7] Gelev S , Gacovski Z , Siea-he X , Yuan-wei J , Deskovski J , "Fuzzy logic in fire control systems fora ir defence", 2007

[8] Shin, J , Hong, M and Hong H , Adaptive Update Rate Target Tracking for Phased Array Radar, IEEE Proc Radar sonar Navig 142 (3): 137p , 1995

[9] Miranda S L C , Baker C J ,Woodbridge K ,Griffiths H D , Fuzzy Logic Approach for Prioritisation of Radar Tasks and Sectors of Surveillance in Multifunction Radar, IET Radar Sonar Navig, 1 (2), 131- 132 p , 2007

[10] Serbest İ , Radar Resource Allocation Optimization In Phased Array Radar Systems, The Graduate School Of Natural And Applied Sciences Of Middle East Technical University ,2019

[11] J Wintenby and V Krishnamurthy, "Hierarchical resource management in adaptive airborne surveillance radars," in IEEE Transactions on Aerospace and Electronic Systems, vol 42, no 2, pp 401-420, April 2006

[12] Bulut M , Tip-2 Bulanık Mantık Sistemlerin Benzetimi İçin Bir Yazılım Geliştirme, Fırat Üniversitesi, Fen Bilimleri Enstitüsü Yüksek Lisans Tezi, 2004

[13] Öztürk M , Paksoy T , Öztürk M , Aralık Tip-2 Bulanık Mantık Tabanlı Sistemlerin Tedarikçi Seçiminde Kullanımının Önemi Üzerine Bir Araştırma, Türkiye Bilişim Vakfı Bilgisayar Bilimleri Ve Mühendisliği Dergisi Cilt:10 - Sayı:2, sf:1-18, 2017

[14] Yıldırım M , Bulanık Mantık Yapay Sinir Ağı ile Doğrusal Olmayan SistemModelleme, Kocaeli Üniversitesi Fen Bilimleri Enstitüsü Yüksek Lisans Tezi, Kocaeli, 1998

[15] Er M , Bulanık sistemler ile faz dizili radarların gerçek zamanlı ayarlanması, Ege Üniversitesi, FBE, Yüksek Lisans Tezi, 2019

A 50 GHz Frequency Selective Surface Design on Double Layer Structure for U Band Applications

O.KASAR[1]

[1] Artvin Çoruh University, Artvin/Turkey, omerkasar@artvin.edu.tr

Abstract - **Frequency Selective Surfaces (FSS), depending on their structure and properties, passes electromagnetic waves at some frequencies while stopping the others at some frequencies. Due to these characteristics, FSSs act as a two-dimensional filter. This filter feature may have band pass, band stop, low pass or high pass characteristics. In this study, a band-pass Frequency Selective Surface was designed for a 50 GHz U Band Applications. The FSS, designed in a two-layer structure. First layer, consist of a double square ring placed in front of the structure. A Jerusalem Cross structure is used on the back layer. The dimensions of the unit cell structure designed for band pass characteristic in the U band were determined using parameter sweep. Then, a 10x10 array was created and the transmission and reflection responses were determined. According to the -3 dB transmission limit, the proposed two-layer U Band FSS is Band-pass between BW = 46.5-53.4 GHz.**

Keywords **- Frequency Selective Surface, U Band FSS, 50 GHz FSS, Double Layer FSS.**

I. INTRODUCTION

FREQUENCY selective surfaces (FSS) show a filter response against electromagnetic waves coming on them [1]. This filter feature may have band pass, band stop, low pass or high pass characteristics. Today, FSSs are widely used in communication and radar applications for filtering the incoming signal while passing through a surface [2]. FSSs are frequently used in applications such as Radome, Satellite communication or wireless communication security [3].

Reflection and transmission characteristics of FSS depend on the type of substrate materials, layer thickness, geometry of metallic patches placed on the layers and periodic placement of the unit cell [1, 4-6]. Today, new radar applications and new generation communication frequencies have started to be made in millimeter wavelengths. This opens the door to new FSS applications [7, 8]. Millimeter wave applications are among the leading FSS applications currently being studied. New FSS unit cell and array designs are made at these frequencies [3, 9].

In this study, a new design was proposed for FSS, which passes the band at U Band frequencies. The unit cell and array with a center frequency of 50 GHz were designed and simulated. In the second part of the study, the properties of the unit cell of the two-layer FSS surface will be mentioned. As an array, $10x10$ dimension of the unit cell has been studied. In the third chapter, simulation results of unit cell and array were given. Success criteria are discussed and bandwidth is specified. All designs were studied in the CST MWS electromagnetic simulation program.

II. U BAND FSS DESIGN

The proposed FSS circuit consists of two layers of surfaces. Designed as a unit cell, FSS is then simulated as a two-dimensional array of cells. FSS, the first layer on the front of the unit cell, consists of doble intertwined square rings. The length of each side of the squares is $b = 2$ and $c = 1.2$ mm, respectively. The line width of each square is $d = 0.3$ mm. There is a gap of g = 0.1 mm between the squares. In Figure 1a, the front view of the upper layer of FSS was given.

On the back surface of the two-layered FSS, the shape known as the "Jerusalem Cross" in the literature was created [1]. In this cross, the line length is e = 2 mm and the line thickness is $f = 0.1$ mm. In Figure 1b, the front view of the back throw was given.

The unit cell placement of two layers formed from square rings and Jerusalem cross structures on the substrate material can be seen from perspective view in figure 1c. It is necessary to use low loss substrate material at millimeter wave and higher frequencies [1]. The substrate material used is ARLON AD300D. The electrical permittivity of the material is $\epsilon_r = 2.94$ tangential loss is $tan\delta = 0.002$ and the material thickness is $h = 0.76$ mm. The copper thickness on the material, which is covered with FSS on one side and is empty at the back, is $t = 0.035$ mm. In Figure 1d, the side view of the two layer FSS was given. As can be seen, it remains as a sandwich between the Jerusalem Cross substrates in the back layer. The width of the unit cell formed was determined by parametric analysis. As will be explained below, the frequency selectivity has been determined by simulating it according to values ranging from $A = 3.1{\sim}4$.

Unit cells come together to form the main large arrays. These arrays can grow as a two-dimensional pattern of unit cells. The larger the pattern, the more the frequency selective surface will cover, and the electromagnetic waves will be selected from the area it covers [1, 10].

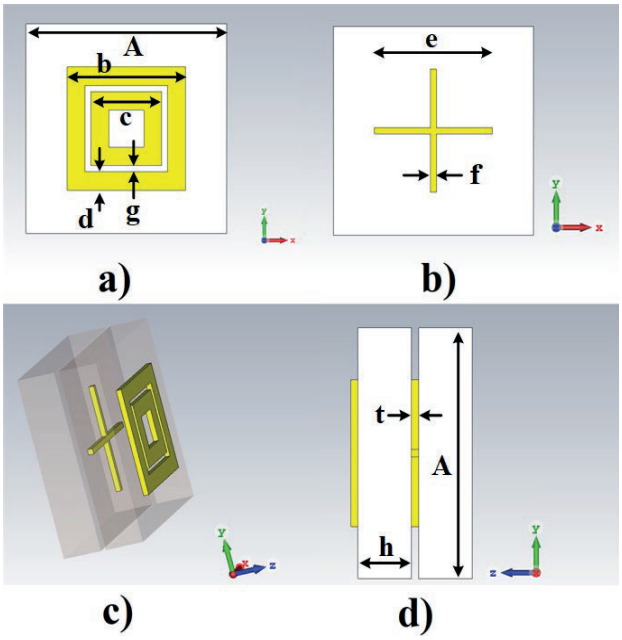


Figure 1: Unit cell Structure of FSS a) front layer, b) back layer, c) perspective, d) side view.

After the unit cell design, a larger scale surface was created to see that the frequency selectivity provided the same characteristic as a surface. With the help of the simulation program, this surface is expected to have the same frequency response [1, 7]. The perspective view of the unit cell in 10 $x10$ scale is given in Figure 2a, and its side view is given in Figure 2b. The dimensions of the structure are square and each length is 10 × A (34 mm).

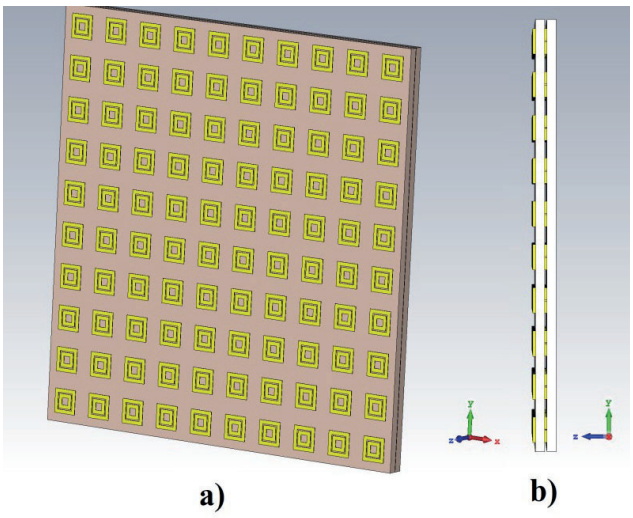


Figure 2: The structure of the array with dimensions of 10x10 is a) perspective, b) side view.

III. Results

The most important criterion for which frequencies a frequency selective surface passes and which does not is to look at the transmission parameter (S_{21}). In the literature, the criterion for accepting FSS in transmission is $S_{21} \geq -3$ dB. In other words, frequency ranges where the transmission parameter is logarithmically greater than -3 dB is considered as pass, while other frequencies are considered as stop. Reflection parameter (S_{11}), another selectivity parameter, is also very important for FSS. At frequencies with high transmission, reflection can be expected to be low.

In the proposed two-layer FSS, the center frequency was designed as $f_c = 50$ GHz. Square rings and Cross shapes were sized according to this frequency. Besides, the outer dimensions of the unit cell, length A, were selected as parametric analysis. Figure 3 shows the response of the transmission graph according to the values of A between 3.1 and 4 mm. At the end of this parametric analysis, $A = 3.4$ mm was selected.

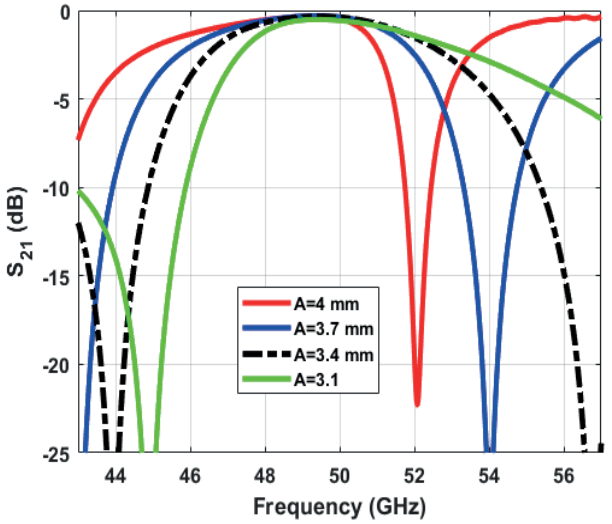


Figure 3: Parametric analysis of unit cell transmission coefficient (S_{21}) vs. frequency plot.

The frequency response of FSS, which grows as a two-dimensional pattern consisting of unit cells, is also expected to be the same as the unit cells [11]. Because the FSS will pass the same frequencies or stop, depend on its size and structure. Therefore, it can be said that the frequency response in the array is somewhat independent of the number of elements. However, during simulation, the size of the array will increase the simulation time since it will affect the number of meshes corresponding to electromagnetic tests solved in numerical methods.

In the proposed design, a 10$x10$ array design was made. Figure 4 shows the response of the unit cell and the array against the frequency of the transmission and reflection coefficients. As can be seen, the frequency selective surface in the proposed structure operates at a frequency of 50 GHz. Its bandwidth is between $BW = 46.5 - 53.4$ GHz.

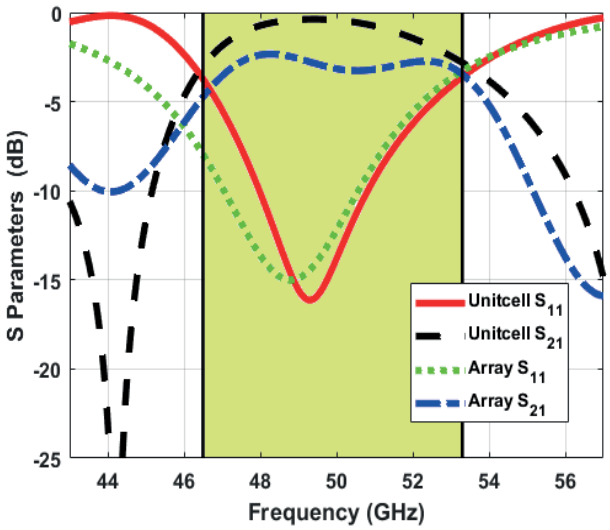


Figure 4: Transmission and reflection graphs of unit cell and array structures.

This proposed design is a successful example of the U Band FSS design as a Band-Pass structure. The fact that its two-layer structure provides the same response for both the unit cell and the array show that the structures selected for FSS can be applied in the U band.

REFERENCES

[1] B. A. Munk, *Finite antenna arrays and FSS*. Wiley, 2003.

[2] Ö. Kasar, M. Geçin, and M. A. Gözel, "Açısal Olarak Değiştirilebilir Dikdörtgen Yamalı Frekans Seçici Yüzeylerle, Ayarlanabilir Bant Geçiren Filtre Tasarımı," *El-Cezeri Journal of Science and Engineering,* vol. 5, no. 3, pp. 756-762, 2018.

[3] S. Çiğdem, C. Tunca, S. Can, and A. E. Yilmaz, "Double bow-tie FSS for X-band and Ku -band operations with its parametric analyses," *Turkish Journal of Electrical Engineering & Computer Sciences,* vol. 24, no. 6, pp. 4751-4762, 2016.

[4] S. Can and A. E. Yılmaz, "Bandwidth enhancement of a triangle with gridded-square loop-loaded FSS for X and K u bands," in *The 8th European Conference on Antennas and Propagation (EuCAP 2014)*, 2014: IEEE, pp. 2150-2153.

[5] F. Güneş, Z. Sharipov, M. A. Belen, and P. Mahouti, "GSM filtering of horn antennas using modified double square frequency selective surface," *International Journal of RF and Microwave Computer-Aided Engineering,* vol. 27, no. 9, pp. 1-8, 2017.

[6] M. A. Belen and P. Mahouti, "Realization of Dielectric Sheets for Gain Improvement of Ultra-Wideband Horn Antennas Using 3D Printer Technology," *Applied Computational Electromagnetics Society Journal,* vol. 34, no. 5, 2019.

[7] K. Mondal, "Bandwidth and gain enhancement of microstrip antenna by frequency selective surface for WLAN, WiMAX applications," *Sadhana,* vol. 44, no. 11, p. 233, 2019.

[8] H.-Y. Chen and Y. Tao, "Performance improvement of a U-slot patch antenna using a dual-band frequency selective surface with modified Jerusalem cross elements," *IEEE transactions on antennas and propagation,* vol. 59, no. 9, pp. 3482-3486, 2011.

[9] W. T. Ramos, R. C. Mesquita, and E. J. Silva, "Frequency selective surface using meander line inclusions," *Journal of Electromagnetic Waves and Applications,* vol. 32, no. 11, pp. 1440-1447, 2018.

[10] P. Mahouti, F. Güneş, M. A. Belen, A. Çalışkan, S. Demirel, and Z. Sharipov, "Horn antennas with enhanced functionalities through the use of frequency selective surfaces," *International Journal of RF and Microwave Computer-Aided Engineering,* vol. 26, no. 4, pp. 287-293, 2016.

[11] M. A. Belen, F. Güneş, P. Mahouti, and A. Belen, "UWB Gain Enhancement of Horn Antennas Using Miniaturized Frequency Selective Surface," *Applied Computational Electromagnetics Society Journal,* vol. 33, no. 9, 2018.

Ball milled Silicon Carbide-Hydroxyapatite Composites for Tetracycline Adsorption

S. BASLAYICI[1], P. BARAN[2], B. GULEN[2] and M. BUGDAYCI[2]

[1] İstanbul Medipol University, Istanbul/Turkey, sbaslayici@medipol.edu.tr
[2]Yalova University, Yalova/Turkey, mehmet.bugdayci@yalova.edu.tr

***Abstract* - The aim of this study is to examine the effects of SiC HA composites synthesized by mechanical alloying method on tetracycline absorption. According to XRD studies, it has been observed that the composite containing 40% hydroxyapatite and 60% silicon carbide successfully absorbs tetracycline. The effects of hydroxyapatite loading amount, solution pH and contact time were investigated by batch experiments. The data obtained are explained with isotherms and kinetic models.**

***Keywords* – SiC-HA, Tetracycline Adsorption, Ball Milling, XRD.**

I. Introduction

Due to the intensive use of antibiotics in treatment applications, it has a high amount of pollutant effect in sea and groundwater. Due to the low absorption of antibiotics into metabolism, a significant amount is exported, which causes the amount of tetracycline in water to be in the range of 0.11–4.20 μg l^{-1}. For this reason, it is important to remove tetracycline from the water[1].

One of the processes used to remove tetracycline from water is adsorption. There are several adsorbents used for tetracycline adsorption such as imprinted polymers, biochar, clays, actived carbon, ceramic composites[2].

Hydroxyapatite finds use in many absorption and biomaterial applications. However, although HA has high biocompatibility, its lack of mechanical strength makes it difficult to use the material alone. Therefore, HA is generally used as a composite by being alloyed with another material. For this reason, composites are made with SiC to improve the mechanical properties of the material while maintaining its positive properties[3].

Various studies have been done to improve the mechanical properties of HA by using sintering additives to consolidate this material during heating. However it is well known that the mechanical properties can also be improved by including a reinforcing phase such as fibers, bristles, platelets or particles in the ceramic matrix[3].

SiC chemical inertness, superior tribological properties, hydroxyapatite-like osseointegration and well-known hemocompatibility make it a very promising candidate for in vitro biosensors, bio-interfaced neural networks, and smart implantable medical devices. Carbon black is used as an absorber in radars due to its superior electrical properties. However, as a single agent, it has several disadvantages such as small absorption peak and narrow absorption bandwidth[4].

On the other hand, SiC is a group IV semiconductor which implies that it is an intrinsic semiconductor which can be n- or p-type depending on the additive used. SiC is often applied in the manufacture of multiband radar absorber materials, as its electrical conductivity can be adjusted[5].

Since the absorbency of SiC as a single absorbent is not satisfactory, it is often used as a co-absorbent in other materials or after surface treatments to increase absorbent capacity[6].

In this study, the effects on tetracycline absorption were investigated by making composites to examine the effects of two materials with superior absorption properties.

II. Materials Method

In the experimental studies, material supplied from Nanokar company with 32 micron size was used as SiC source. The hydroxyapatite used as an additive material is Xi'an Realin Biotechnology Co., Ltd. and its purity is over 99%.

Mechanical alloying method has been used to make silicon carbide and hydroxyapatite composites. The stainless steel balls used during this process. As the process progresses, the surface area of the components that are welded and disintegrated increases and it is possible to bond with the activated surfaces. As a result of the mechanical alloying process, the two materials are obtained as composite. Mechanical alloying device presented in Figure 1.

Figure 1: Planetary Ball Mill.

In the study, after the SiC and HA powders were weighed on a sensitive scale, they were dried for 2 hours at 100 ˚C in furnace, and then loaded to the planetary ball mill. In the experiments, the SiC / HA ratio was added to 90/10, 80/20,

70/30 and 60/40 respectively, and the ball/powder ratio was taken as 1/10. Experiments performed at 380rpm rotational speed and 6 hours.

BrukerTM D8 Advanced Series X-Ray diffractometer was used to monitor the changes in the structure of the powders in characterization studies. While performing the analyzes, the 10-90 range in CuKα (k = 1.5406 Å) radiation was scanned at 2 degrees / minute and the interpretation procedures were performed with the Xpert High Score program. The microstructures of the samples were examined using the Scanning Electron Microscope (SEM) technique, and the studies were carried out with the Zeiss Gemini 500 NeoScope device.

III. Results and Discussion

In experimental studies, firstly, tetracycline absorption capabilities of SiC and HA were investigated separately and it was determined that both materials have limited absorption ability.

Then, SiC / HA ratio was changed as 90/10, 80/20, 70/30 and 60/40 and in the experiment performed at 60/40 ratio in accordance with the change ratio, a structure containing more of both components was obtained. Accordingly, the absorption ability of HA added at 60/40 ratio was analyzed by XRD technique and the results are given in Figure 2.

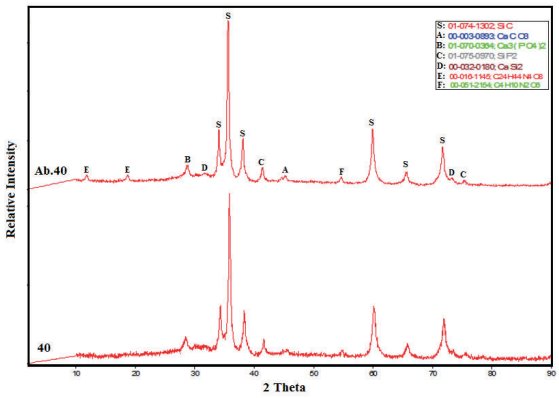


Figure 2: XRD pattern of SiC/HA 60/40 composites.(Ab 40 absorbed, 40 ball milled).

The E and F phases seen in Figure 4 are not seen in the test peaks (40) performed by mechanical alloying and are seen in the graph obtained after absorption (Ab.40). This proves that the composite structure has successfully absorbed tetracycline.

After XRD analysis, SEM analysis was applied to observe the changes in the microstructure of the samples, and the results are presented in Figures 3 and 4.

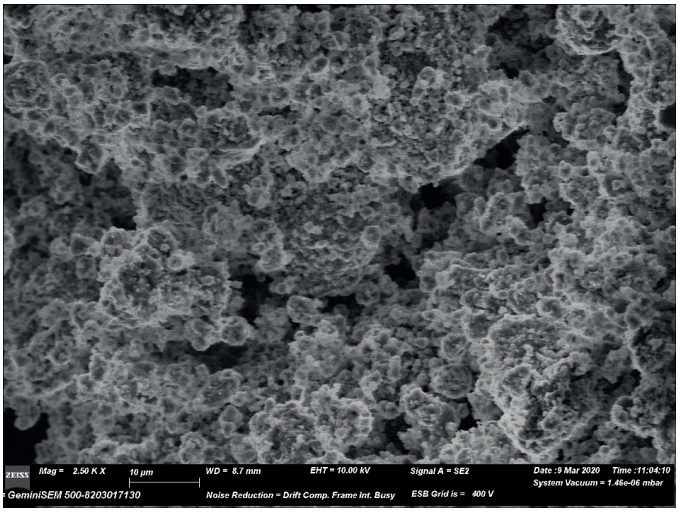


Figure 3: SEM image of SiC/HA 2500 X.

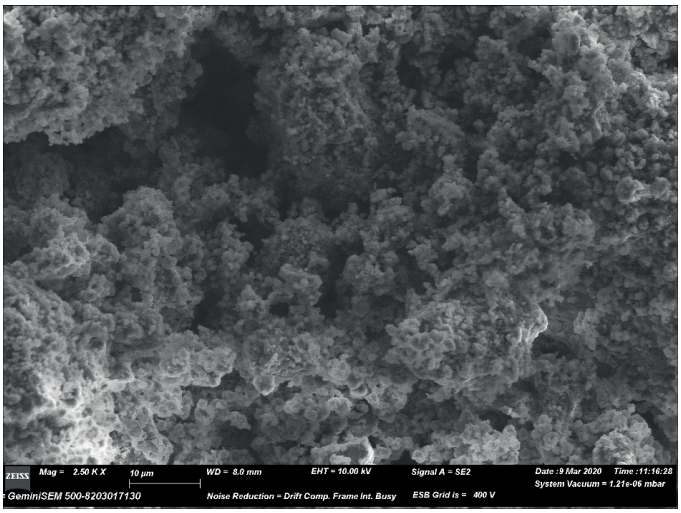


Figure 4: SEM image of SiC/HA tetracycline absorbed 2500 X.

When Figure 3 and Figure 4 are compared, it is seen that the material has become more compact due to absorption. Figure 3 powders more porous than Figure 4. Because Figure 4 absorbed tetracycline and became more dense than Figure 3.

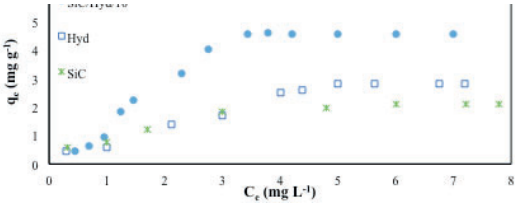


Figure 5: Adsorption isotherms of raw SiC and hydroxyappatite materials.

Fig. 5 shows the adsorption isotherms of raw SiC and hydroxyapatite materials with comparison SiC/HA composite. Significant increase of tetracycline adsorption was observed using SiC/HA composite. Adsorption capacities were achieved as 1.95, 2.80 and 4.55 mg g-1 for SiC, HA and SiC/HA/10, respectively. The reason of increase in adsorption was the synergetic effect of SiC and HA materials, which resulted with the combination of adsorption sites to adsorb tetracycline in aqueous solution.

IV. Conclusion

In conclusion, SiC / HA composite material has been successfully produced by mechanical alloying method. The effects of the produced material on tetracycline absorption were examined and it was observed that the composite material successfully absorbed.

References

[1] P Demircivi, M Bugdayci, Synthesis and thermodynamic simulation of Zr Doped SrAl2O4 ceramic powders: Enhanced adsorptive properties for antibiotic tetracycline, Mater Res Express 6 (2019) 86506 https://doi org/10 1088/2053-1591/ab1ae4

[2] C Ian, R Marilyn, Tetracycline Antibiotics: Mode of Action, Applications, Molecular Biology, and Epidemiology of Bacterial Resistance, Microbiol Mol Biol Rev 65 (2001) 232–260 https://doi org/10 1128/MMBR 65 2 232

[3] S Hesaraki, T Ebadzadeh, S Ahmadzadeh-Asl, Nanosilicon carbide/hydroxyapatite nanocomposites: Structural, mechanical and in vitro cellular properties, J Mater Sci Mater Med 21 (2010) 2141–2149 https://doi org/10 1007/s10856-010-4068-7

[4] X Liu, Z Zhang, Y Wu, Absorption properties of carbon black/silicon carbide microwave absorbers, Compos Part B Eng 42 (2011) 326–329 https://doi org/10 1016/j compositesb 2010 11 009

[5] Z A Yaşar, R A Haber, Effect of Carbon Addition and Mixture Method on the Microstructure and Mechanical Properties of Silicon Carbide, Materials (Basel) 13 (2020) 3768 https://doi org/10 3390/ma13173768

[6] K Guan, L Zhang, F Zhu, H Li, H Sheng, Y Guo, Multi-layer SiC-graphene oxide-hydroxyapatite bioactive coating for carbon/carbon composites, J Alloys Compd 821 (2020) 153543 https://doi org/10 1016/j jallcom 2019 153543

Production of Vanadium Carbide via Magnesiothermic Reduction

M. BUGDAYCI[1] and A.TURAN[1]

[1]Yalova University Chemical Engineering Dept., Yalova/Turkey, mehmet.bugdayci@yalova.edu.tr

Abstract **- Vanadium carbide is important for industrial applications because of its high temperature resistance, high chemical and thermal stability. It is generally obtained from the reaction between V and C powders at a high temperature ranging from 1100 to 1500◦C. Investigations on these high strength, high abrasion resistant, hard materials have been intensified in recent years and consequently significant improvements have been achieved. In this study, VC alloys produced by high cost processes will be produced by reducing the oxides of their components by Self Propagating High Temperature Synthesis methods. V_2O_5 will be used as oxidized Vanadium source, C_{black} as carbon source, magnesium as reductant will be used. Furthermore, the effect of different stoichiometric charge components will be realized by XRD, XRF, AAS and optical microscope analyzes for different reductants.**

Keywords **– SHS, Ball Mill, VC, Carbides.**

I. Introduction

Carbides and nitrides of the transition metals of the fourth to sixth group of the periodic table combine the classical refractory properties with similar physical properties of metals. Interest in these physical properties is increasing due to the electrical-thermal conductivity, Hall coefficient and magnetic susceptibility of these alloys. Therefore, carbides made by transition elements have become promising candidates for their use as heterogeneous catalysts [1]. Hydro-treatment of hydrocarbon raw materials, including VC and VN, with transition metal carbides and nitrides is an example of their use [2]. Decontamination of fuels from harmful components has become an issue of great importance due to increasingly stringent environmental regulations on the composition of fuels used in transportation [3]. Vanadium carbide (VC) acts as a catalyst in such applications due to its high surface area properties, and is especially used in the purification of sulphurous compounds [4-6].

In metallurgical applications, grain refining is an important process used to increase strength. Small concentrations (0-01%) of strong carbide-forming elements such as niobium, titanium and vanadium are added to mild steels to raise ferritic steels to higher strength levels by grain refining. The resulting fine-grained ferrite structure is usually further strengthened by a fine dispersion of a particular alloy carbide. Due to the nature and origin of these alloy carbide dispersions, it is considered a popular field of study. The alloy carbides added to ferritic steel restrict the dislocation movements by settling in the precipitated grains, as a result of this limitation, dislocations are locked and act as a frank-read source and encourage the formation of new fine grains. The newly formed fine-grained structure enables the material to reach higher strength values and to be used in more complex applications [7,8].

Studies on both the friction coefficient and wear resistance of vanadium carbide have shown that the material has a good tribological performance. Due to these properties, the material is used in coating applications to protect backing materials operating under high wear conditions [9]. Transition metal carbides, due to their superior properties such as high hardness, high melting point, high thermal conductivity, high strength even at high temperatures, high wear-corrosion resistance and high chemical stability, such as cutting tools, catalyst applications, wear-resistant parts and high temperature building materials has found extensive applications in many areas [10-12].

II. Materials Method

Self-propagating high temperature synthesis (SHS) is a simple and low cost method used in the production of high-tech ceramics and intermetallic materials. The operation of the method is that the excessively exothermic reaction starts spontaneously with the trigger and proceeds spontaneously over the reaction mixture in the form of a wave. For these events to occur, the reaction must have a relatively high activation energy and generate extremely high heat. In the SHS method, after the first heat is given to the starting mixture to bring the mixture to the ignition temperature, the reaction proceeds automatically without the need for any external energy. The heat released in each layer passes to the next unreacted mixture layer and creates the first heat for the zone to ignite and raises the temperature. While the front of the SHS reaction leaves the combustion products behind, it moves towards the unreacted mixture and separates the heat-affected zone and the reaction zone. The high amount of heat energy generated during the application directly affects the reaction rate; It provides a very economical and highly efficient production with increasing speed. SHS process schematic image presented at Figure 1.

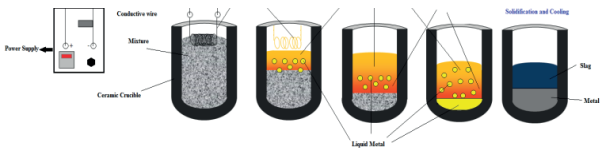


Figure 1: SHS Process.

There are some advantages in this technique such as, less

energy requirement, very short processing time, simple process, sub-micron sized particles, high purity products. Although there are many advantages, this technique have some defects. These are controlling reactions, cost of reductants, and controlling quality of final products. For this reason thermodynamic data's "Specific heat and adiabatic temperature" have very important role for predict reaction circumstances.

Specific heat; determining the heat evolved during the reaction and to estimate if the temperature achieved is sufficient to smelt the metal and the slag and to separate them due to the different density.

Calculated by dividing the enthalpy of the reaction at 25 °C by the sum of the molecular weights(MW) of the reaction products.

$Me_IO + Me_{II} \rightarrow Me_I + Me_{II}O \quad \Delta G_R = \Delta GMe_{II}O - \Delta GMe_IO < 0$

Specific heat < 2250 J/g, insufficient to melt the charge and to separate metal/slag

Specific heat > 4500 J/g, violent reaction and may even be explosive.

Specific heat : 2250 – 4500 J/g, a controlled and self-sustaining metallothermic reaction

Adiabatic Temperature must be higher than 1527 °C to start metallothermic reaction.This study aims to produce Iron based alloy from mill scale via aluminothermic method and investigate effect of Si and Mg addition.

In the experiments, the powders were weighed in the determined stoichiometric proportions on precision scales and dried at 105ºC for 30 minutes After the charge was mixed homogeneously and loaded into the crucible, the reaction was started with the help of variac and the reduction was occur automatically. At the end of the experiment, final products were weighed, analyzed by XRD, AAS methods.

The reaction stoichiometry was calculated according to Equation 1. This amount of Mg was chosen as a 100% stoichiometric ratio. Then, according to this ratio, the other mixtures were calculated as 95%, 105%, 110% by changing the amount of Mg.

$$V_2O_5+2C+5Mg \rightarrow 2VC+5MgO \quad (1)$$

III. Results and Discussion

After the reactant materials were reduced with the SHS method, the products were weighed and subjected to XRD analysis. The analysis result of the sample containing 95% Mg as stoichiometric is given in Figure 2.

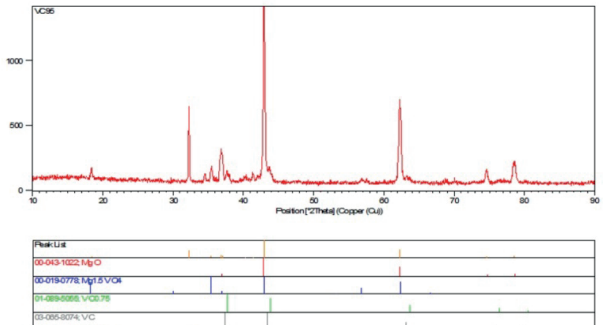


Figure 2: XRD Result of %95 Stoichiometric Samples.

When Figure 2 is examined, it is seen that the main phase formed as a result of the reaction is MgO. In addition to this phase, it was seen that Mg bonds with vanadium and the 2nd active phase was formed in this way. When Figure 2 is examined, it is seen that the VC phase is obtained at 38, 42 and 78 degrees in addition to the mentioned phases. The XRD analysis results of the 100% stoichiometric mixture obtained in the next stage of the study are given in Figure 3.

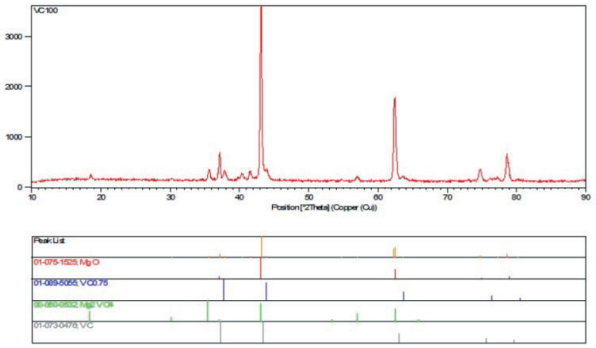


Figure 3: XRD Result of %100 Stoichiometric Samples.

According to Figure 3, it was observed that the effect of the MgVO phase shown in Figure 1 decreased and more VC was formed. It has been observed that the increase in Mg stoichiometry can reduce more Vanadium Penta Oxide by increasing the reaction energy.

The samples for which the magnesium stoichiometry was determined as 105% and 110% were analyzed by XRD method, and the findings are given in Figure 4 and Figure 5, respectively.

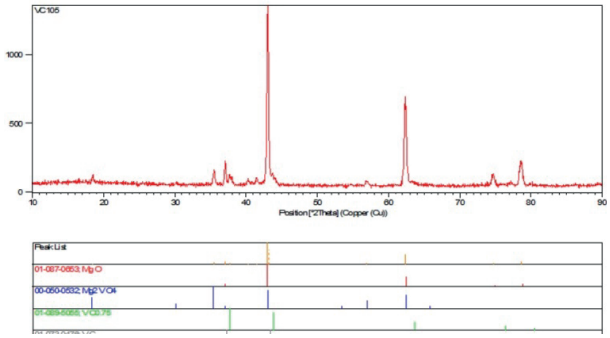


Figure 4: XRD Result of %105 Stoichiometric Samples.

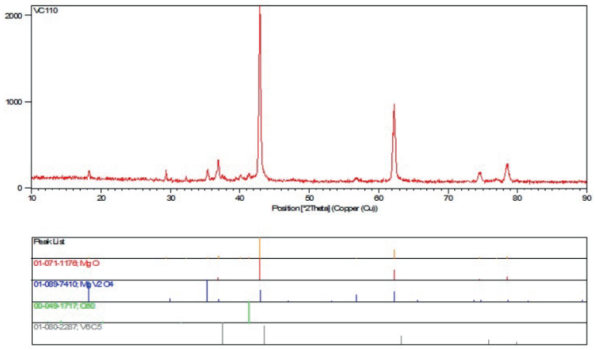


Figure 5: XRD Result of %110Stoichiometric Samples.

According to Figure 4 and Figure 5, the increase in Mg stoichiometry increased the amount of Mg in the medium and caused an increase in MgV based phases. Accordingly, 100% stoichiometric composition was determined as the mixture in which VC was obtained under the most optimum conditions.

According to the XRD results, the VC phase was obtained at each stoichiometry. However, in order to use the product, it must be purified from Mg based impurities. This work is an ongoing process. In the next stage of the study, it is aimed to obtain a product with XRD peaks equivalent to the commercial quality product by removing impurities by leaching method.

IV. Conclusion

In conclusion, VC was obtained from oxidized raw materials by SHS method. However, when the XRD analysis of the samples were examined, it was determined that there were magnesium-based pollutants in the structure besides VC. Leaching was planned to be carried out in the later stages of the study to eliminate these impurities.

Acknowledgment

The authors would like to thank the Yalova University BAP Project “2019 / AP / 0024” for its financial support.

References

[1] Ramanathan, S and Oyama, S T , New catalysts for hydroprocessing: transition metal carbides and nitrides J Phys Chem , 1995, 99, 16365±16372

[2] H Preiss, D Schultze and K Szulzewsky Carbothermal Synthesis of Vanadium and Chromium Carbides from Solution-Derived Precursors Journal of the European Ceramic Society 19 (1999) 187±194

[3] P Rodrıguez , J L Brito, A Albornoz, M Labadı , C Pfaff , S Marrero , D Moronta , P Betancourt Comparison of vanadium carbide and nitride catalysts for hydrotreating Catalysis Communications 5 (2004) 79–82

[4] C Guillard, M Lacroix, M Vrinat, M Breyse, B Mocaer, J Grimblot, T Des Courieres, D Faure, Catal Today 7 (1990) 587

[5] C E Scott, B P Embaid, M A Luis, F Gonzalez-Jimenez, L Gengembre, R Hubaut, J Grimblot, Bull Soc Chim Belg 104 (1995) 331

[6] M Magarelli, C E Scott, D Moronta, P Betancourt, Catal Today 78 (2002) 339

[7] A T Davenport, F G Berry, and R W K Honeycombe, Metal Sci J , 1968, 2, 104

[8] J M Gray and R B G Yeo, Trans Amer Soc Metals, 1968, 61,255

[9] Y Herrera, I C Grigorescu, J Ramirez, C Di Rauso, M H Staia Microstructural characterization of vanadium carbide laser clad coatings

[10] Stroms EK The refractory carbides, refractory materials series New York: Academic Press; 1967

[11] Toth LE Transition metal carbides and nitrides New York: Academic Press; 1971

[12] Pierson HO Handbook of refractory carbides and nitrides Westwood: Noyes Publications; 1996

Fabrication of copper centered metal organic framework and nitrogen, sulfur dual doped graphene oxide composite as a novel electrocatalyst for oxygen reduction reaction

SEYED ALI MOUSAVI[1] , MEHDI MEHRPOOYA[1] , FATHOLLAH POURFAYAZ[1]

[1] Department of renewable Energies and Environment, Faculty of New Sciences and Technologies, University of Tehran,Tehran, Iran, s.a mousavi74@ut.ac.ir

[1] Department of renewable Energies and Environment, Faculty of New Sciences and Technologies, University of Tehran,Tehran, Iran, mehrpoya@ut.ac.ir

[1] Department of renewable Energies and Environment, Faculty of New Sciences and Technologies, University of Tehran,Tehran, Iran, _____

***Abstract* - Today, the most common catalyst used for Oxygen Reduction Reaction (ORR) at the cathode side of fuel cells, is the rare and expensive platinum metal. The main focus of this study is to synthesize a free – platinum electrocatalyst for ORR applications. For this purpose, the Copper - centered Metal Organic Framework (Cu - MOF) is selected as electrocatalyst. The electron conductivity of MOFs is low. So, in order to enhance the ORR kinetics and conductivity, for the first time, Nitrogen and Sulfur Dual-Doped Reduced Graphene Oxide (NS - RGO) with different concentrations are incorporated into the Cu – MOF structure. For evaluating the structural properties and morphology of synthesized electrocatalysts, three important characterization techniques are employed. Also, in order to assess the durability and ORR activity, the electrochemical measurements are performed. The electrochemical tests are implemented by using the Rotary Disk Electrode (RDE) device. The best ORR activity is related to the 8% NS – RGO - Cu – MOF. The onset potential and electron transferred number (n) of this catalyst are obtained to be -0.06 V vs Ag/AgCl and 3.53, respectively. In other words, it tends to favor the 4e- pathway for oxygen reduction reaction. Finally, the electrochemistry activity of synthesized electrocatalysts is compared to the previous investigations and commercial 20 wt% Pt/C. It can be concluded that the incorporation of carbon-based materials and MOFs has an outstanding ORR performance, and can be replaced with the platinum-based catalysts.**

Keywords* - *Oxygen Reduction Reaction; MOF; Reduced Graphene Oxide; Fuel Cell; Electrocatalyst;

I. INTRODUCTION

The high cost of the catalysts used at the anode and cathode sides limits the commercialization of fuel cells for practical applications.

One of the effective factors on performance of fuel cells and batteries is the type of catalysts used at the anode and cathode sides. At the moment, the most common catalyst for ORR is platinum (Pt) or platinum group metals [1]. Platinum element is very costly and rare, so in order to speed up to commercialization of fuel cell technology, the platinum – free catalysts must be replaced with platinum [2]. The electrochemical reaction, which takes place at the cathode side is the Oxygen Reduction Reaction (ORR). The sluggish kinetics of ORR is one of the main challenges in fuel cell systems [3]. In order to diminish the polarization loss over the ORR, many researches have been carried out to develop the lasting and effective ORR catalysts consist of carbon based materials, metals, metal oxides, and metal – organic complexes [4].

Metal organic frameworks (MOFs) represent a class of porous material which is formed by strong bonds between metal ions and organic linkers [5]. MOFs due to their crystal and three-dimensional structure have no an acceptable electron conductivity. High electrical conductivity is one of the most basic and important properties for any electrocatalysts used in fuel cells. In other words, the low value of electrical conductivity of MOFs has been limited the use of them as stand – alone for catalyst [6]. So, in order to enhance the electrical conductivity of MOFs, carbon-based materials such as, Carbon Nano Tube (CNT) and Graphene Oxide (GO), are incorporated to the MOF. In following, the recent investigations on using MOFs for ORR applications are expressed. To evaluate the structure and morphology of each electrocatalyst, the physicochemical characterization tests, including X-ray diffraction (XRD) pattern, Raman spectroscopy, Scanning Electron Microscopy (SEM), Fourier transformed infrared (FTIR), and Energy-Dispersive X-ray Spectroscopy (EDS) are employed. Also, in order to assess the electrochemistry activity and durability of catalyst, the electrochemical measurements, including Cyclic Voltammetry (CV), Linear Sweep Voltammetry (LSV), and chronoamperometric response are used. So, in investigations of catalyst, the performance of electrocatalyst is evaluated by the physicochemical characterization and electrochemical tests.

Fe - centered metal MOF is fabricated for ORR in an alkaline electrolyte [7]. According to the electrochemical test, by applying this catalyst, the ORR occurs by the four-electron pathway. Also, the onset potential is computed to be -0.12 V vs Ag/AgCl. It was found that MOFs can open a novel rout for fabrication the catalysts with low cost and high efficiency. The bifunctional catalyst of Co/MIL101(Cr) – O is synthesized and evaluated [8]. According to the electrochemical test, the onset

potential is reported to be -0.12 V vs Ag/AgCl. In other research, a mixed metal MOF (Fe/Co – MOF), which its organic ligands is trimesic acid is synthesized and introduced for ORR in lithium air batteries [9]. The onset potential and electron transferred number (n) are found to be -0.13 V and 3, respectively. In order to enhance the electrochemistry activity for ORR, Co-MOF is synthesized on carbon as the support [10]. The electrochemical measurements are performed in alkaline medium, and the onset potential is gained to be 0.8 V vs. RHE. In addition, the value of n is accounted to be 3.4 at 0.7 V vs. RHE. The hybrid Co-OBA and carbon black is fabricated, and assessed from the morphology and activity viewpoints [11]. Benzoic acid is utilized as the organic ligands. The outcomes show that the synthesized catalyst has a suitable lasting in alkaline medium, and the onset potential is obtained to be -0.197 V vs Ag/AgCl. A novel assembly of Co - centered MOF and MWCNT (Co-MOF@CNTs 3D) is fabricated and evaluated for ORR [12]. Because of the electrostatic interaction, Co^{2+} ions are adsorbed onto the CNTs. The onset potentials of CNT, Co – MOF, and Co-MOF@CNT (5%) are found to be 0.85, 0.82 and 0.91 V vs RHE, respectively. Also, the electron transferred number is computed by 3.6. As a result, CNT can improve the poor electron conductivity of MOFs. Jahan et al. [13] fabricated a novel GO and copper - centered metal MOF composite electrocatalyst for ORR. The electrochemical measurements are implemented in acid media, and it was found that the durability of synthesized catalyst is higher than pure MOF. Also, the power density is found to be 76% of Pt catalyst. The new composite of $Co@Co_9S_8$-N/C as a non-noble-metal catalyst for ORR is fabricated and analyzed [14]. The results of electrochemical tests show this catalyst has an onset potential of 0.999 V vs. RHE and an ORR peak potential of 0.779 V vs. RHE. In fact, MOFs can help to develop the fuel cell technology to commercialization.

In present investigation, for the first time, in order to improve the electrical conductivity of copper – centered MOF (Cu – MOF), NS – RGO is hybridized with this type of MOF. First, GO is fabricated by modified Hummer's method. Next, the synthesized GO is reduced by thiourea, and N and S are doped on GO sheets (NS – RGO). In continue, the different concentrations (X%) of NS – RGO are added to the MOF structure to produce the final catalyst powder (X% NS – RGO – Cu – MOF). The electrocatalysts are fabricated for employing in ORR of cathode side. The morphology and kinetic of synthesized electrocatalysts are assessed by physical and electrochemical measurements, respectively. The electrochemical tests are implemented in alkaline media. Finally, the ORR activity of synthesized electrocatalysts is compared to the 20 wt% Pt/C catalyst. In fact, the main focus of this investigation is to synthesize a new electrocatalyst with a simple hydrothermal method, which can be a suitable and affordable alternative to the platinum (Pt) for ORR applications.

II. METHODOLOGY

All of chemical materials have been purchased from the Merck with analytical grade. Also, the purchased materials were utilized without further purification. Graphite powder and Nafion solution (5 wt%) were provided from the Sigma-Aldrich.

In this investigation, GO was synthesized based on the modified Hummer's method. For reduction of GO and dual doping of nitrogen and sulfur (NS – RGO), thiourea (CH_4N_2S) was utilized. In this method, first, 50 mg of GO was dispersed in a 50 mL of ethylene glycol ($C_2H_6O_2$) (EG) by using the ultrasonic bath device for 1 h. In following, 0.5 g of thiourea was dispersed in a 78 mL of EG, and thiourea solution was achieved. Next, the thiourea solution was added to the GO dispersion, and sonicated by applying the probe sonicator for 90 minutes. Then, the solution was transferred to a round bottom flask and exposed to the reflux for 5 h at 190 °C. After completion of this hydrothermal process, the NS – RGO powder was collected by centrifugation and thoroughly washed with DI water and ethanol. The solution was dried at 80°C in the vacuum oven overnight.

The main innovation point of this investigation is adding the different quantities of GO and NS – RGO to the MOF mixture. In this method, the weight fraction of 6 and 8% of GO and NS – RGO were added to the mixture, and sonicated by the probe sonicator for 1h. In following, the homogenized mixture was transferred to a 50 mL autoclave and then heated at 120 °C for 36 hours. After the washing by DMF, the mixture was transferred to a vacuum oven. Due to the adding NS – RGO to the mixture, the color of dried powder was changed from turquoise blue to blue grayish. In order to achieve a conductive and active MOF, the carbonization method is employed [15]. Therefore, the 8% NS - RGO - Cu- MOF powder was transferred to a tubular furnace for thermal annealing at 750 ° C in a nitrogen (N_2) atmosphere.

Figure 1 displays the schematic of the chemical structures of synthesized electrocatalysts. As can be shown, graphene structure lacks oxygen functional groups. After oxidization of graphene, oxygen functional groups are formed on carbon atoms, which including epoxy, carbonyl, carboxyl, and hydroxyl. In the Reduced Graphene Oxide (NS -RGO) structure, the amount of oxygen functional groups has been considerably decreased compared to GO, and the structure is closer together. Cu – MOF has a 3D framework, which consists of two different ligands: $C_8H_6O_4$ & $C_6H_{12}N_2$. The Cu – MOFs are formed on GO and NS – RGO sheets.

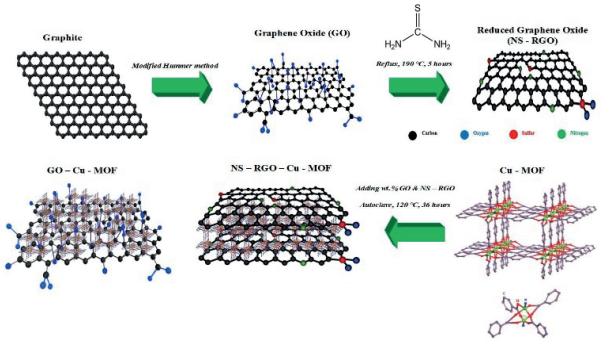


Figure 1. The schematic of the chemical structures of synthesized electrocatalysts.

The performed physical tests in this work include X-ray diffraction (XRD) pattern and Scanning Electron Microscopy (SEM). Also, the electrochemical measurements are implemented by applying the three-electrode cell system. In order to prepare the catalysts ink, 2 mg of catalyst powder were mixed with 1 mL of the water and alcohol solution (1:1), and 10 mL of 5 wt% Nafion solution. After that, the mixture was sonicated in an ultrasonic bath device for 3 hours.
At the final stage of electrochemical studies, the obtained results were compared to the previous studies and 20 wt% Pt/C commercial catalyst.

III. RESULTS AND DISCUSSIONS

Figure 2 indicates the XRD patterns of synthesized electrocatalysts. In XRD pattern of GO, a sharp peak (001) at $2\theta = 11°$ is shown. Also, the intensity of this peak is about 2800. This peak with a d – spacing of 0.8 nm is related to the GO, and is similar to the reported results by Jabbar el al. [16] as reference paper. As can be seen in XRD pattern of NS – RGO, after reduction of GO by thiourea, the peak associated with GO (001) is disappeared. A new peak at $2\theta = 25°$ is appeared (002), and d – spacing is 0.4 nm. It can be confirmed that N and S agents have replaced to the oxygen functional groups (similar to [17]). The XRD patterns show that new peaks are appeared, which it indicates that Cu – MOF seeds are formed on GO and NS – RGO surface. Also, two major peaks associated with Cu – MOF are observed in XRD patterns of 6% GO – Cu – MOF and 6% NS - RGO – Cu – MOF catalysts. But, the intensity of these peaks has been reduced, which it illustrates that the crystallite of MOF has been decreased. In XRD pattern of 6% GO – Cu – MOF (D), the peak associated with GO (001) is shown at $2\theta = 11.64°$, but due to the formation of Cu – MOF on GO sheets, the intensity of this peek reduces considerably. The intensity of NS – RGO (002) has been reduced. For example, the peak at $2\theta = 13.5°$ is related to the pure copper. Figure 5 (F) shows the XRD pattern of 8% NS - RGO – Cu – MOF after the thermal annealing process. each Cu^{2+} is coordinated by 4 oxygen atoms. Actually, after the carbonization process, the crystalline structure of Cu – MOF is destroyed, and active sites for ORR increases. The peak (111) at $2\theta = 36.29°$ is related to the RGO – Cu_2O composite [18].

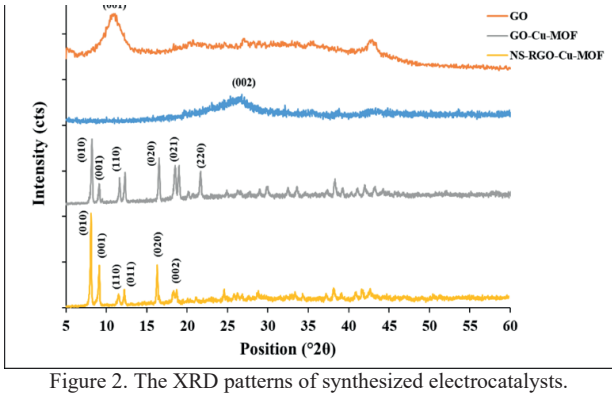


Figure 2. The XRD patterns of synthesized electrocatalysts.

Figure 3 shows the SEM images of synthesized electrocatalysts. The sheets and flakes of GO have a uniform structure. The shape of GO sheets is wavy and folded. The morphology of synthesized GO is similar to the previous investigations [19, 20]. As a result, GO powder has been synthesized properly. As can be seen in SEM image of NS – RGO, the structure of reduced graphene oxide is similar to graphene oxide. In NS-RGO, due to the high temperature reflux heat treatment (180 ° C), and ultrasonic operation, the sheets have become smaller and closer to each other. Actually, by replacing the N and S agents with the oxygen functional groups, the interaction between the sheets has been increased, which it is the clearest difference between the structure of NS-RGO and GO. The morphology of NS – RGO sheets are confirmed with the SEM images reported by Bag et al. [19]. The SEM image of Cu – MOF are exhibited in part C. It can be seen that the crystal structure of MOF has been formed properly. The size of crystals is uniform, and is in the nanometer range (333.3 nm, 480.2 nm, 168.9 nm). As can be seen in SEM images of 6% GO - Cu – MOF, the MOF crystals are formed on GO sheets. Also, the distribution of Cu – MOF particles on GO is uniform. According to the SEM images of 6% and 8% NS – RGO - Cu – MOF, the more amounts of MOF crystals have been formed on NS – RGO sheets in comparison to 6% GO – Cu – MOF. This is due to the presence of sulfur and nitrogen agents, which act as an anchor, leading to a strong bonding of the metal ions to the carbon sheets. Metal ions are more easily absorbed by more negative charge points. The electronegativity of sulfur and nitrogen is higher than carbon, thereby causing the negative charge and pulling metal ions (Cu^{2+}) to the surface.

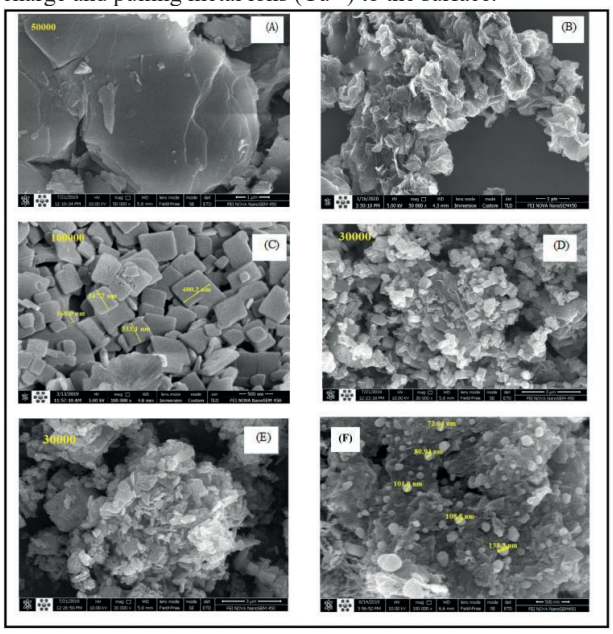


Figure 3. The SEM and TEM images of synthesized electrocatalysts: (A) GO; (B) NS – RGO; (C) Cu – MOF; (D) 6% GO - Cu – MOF; (E) 6% NS – RGO - Cu – MOF; (F) 8% NS – RGO - Cu – MOF.

Figure 4 depicts the CV curves of synthesized electrocatalysts. In all catalysts, the peak associated with ORR is shown. Also, in presence of N_2 gas (inert gas) no peak is appeared. It is concluded that, the synthesized electrocatalysts have an acceptable selectivity property. For GO catalyst, the ORR peak is at the -0.3 V vs Ag/Ag/Cl, and the current peak is found to be -0.7 mA/cm^2. This peak is related to the ORR, and can be

confirmed by [19]. For NS – RGO, the ORR peak is appeared at -0.24 V vs Ag/Ag/Cl, and the current density is obtained to be -1.8 mA/cm^2. With the reduction of GO, the ORR peak has become sharper, and shifted to the more positive values, which it indicates the electrochemistry activity is improved. The ORR peak of Cu – MOF is at -0.43 V vs Ag/Ag/Cl. To enhance the electron conductivity and electrochemistry activity, the Cu - MOF are hybridized with GO and NS – RGO. As can be noticed, the ORR peaks for 6% GO - Cu – MOF, 6% NS – RGO - Cu – MOF, and 8% NS – RGO - Cu – MOF are shifted to -0.41, -0.36, and -0.08 V vs Ag/Ag/Cl, respectively.

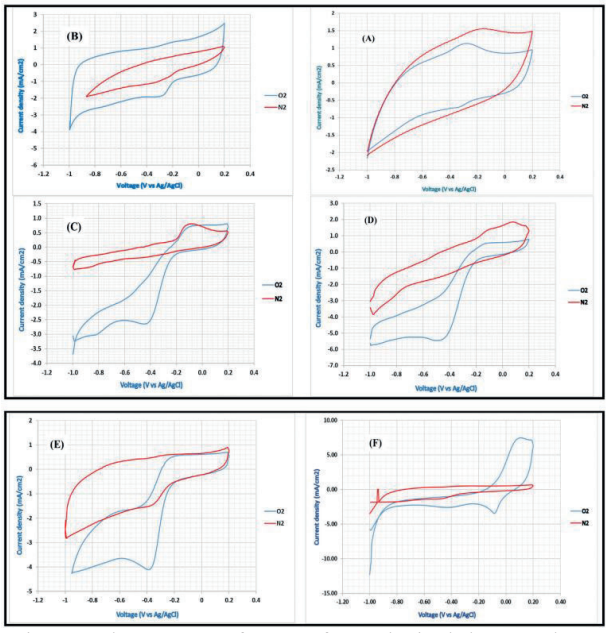


Figure 4. The outcomes of CV test for synthesized electrocatalysts: (A) GO; (B) NS – RGO; (C) Cu – MOF; (D) 6% GO - Cu – MOF; (E) 6% NS – RGO - Cu – MOF; (F) 8% NS – RGO - Cu – MOF.

Figure 5 displays the outcomes of LSV test for each synthesized electrocatalyst, and 20 wt% Pt/C commercial catalyst. By applying the LSV test, the value of onset potential is achieved. As can be observed, in all curves, by increasing the rotation speed of the working electrode, the oxygen mass transfer resistance decreases and the oxygen movement towards the electrocatalyst becomes smoother. In other words, with the increase of ω, the amount of oxygen on the electrocatalyst surface increases, resulting in more current density. The onset potential for GO is about -0.16 V vs Ag/Ag/Cl. Also, the current density is obtained to be -1.5 mA/cm^2 at -1V vs Ag/Ag/Cl. By reducing the GO with thiourea (NS – RGO), the onset potential of catalyst was shifted to -0.1 V vs Ag/Ag/Cl. Actually, by doping the nitrogen and sulfur agents on GO structure, the kinetic of electrocatalyst has been improved. According to the polarization curves, the current density of catalyst enhances. For instance, the current density of catalyst increases by 180% (-4.2 mA/cm^2) in comparison to the GO at -0.1 V vs Ag/Ag/Cl and 2500 rpm. Based on the polarization curve of Cu – MOF, the onset potential is about -0.22 V vs Ag/Ag/Cl. Also, the values of current density at different rotation speeds are low. Correspondingly, to improve the poor electron-conductive of Cu - MOF, the different percentage of GO and NS – RGO were incorporated to the MOF structure. The onset potential of 6% GO - Cu – MOF has been shifted from -0.22 to -0.11 V vs Ag/Ag/Cl. By adding 6 and 8%wt of NS – RGO, the onset potentials are found to be -0.08 and -0.06 V vs Ag/Ag/Cl, respectively. It can be concluded that the combination of carbon-based materials with MOFs significantly improves the kinetics and activity of catalysts in ORR applications. In Figure 5 (H), the LSV curves of synthesized electrocatalysts at 2500 rpm are compared to the 20 wt% Pt/C. The onset potential of 20 wt% Pt/C is found to be about zero. Amongst the synthesized electrocatalysts, 8%wt of NS – RGO has the best performance and electrochemistry activity in the ORR. Also, it is suggested that other concentrations of NS – RGO are evaluated in future works.

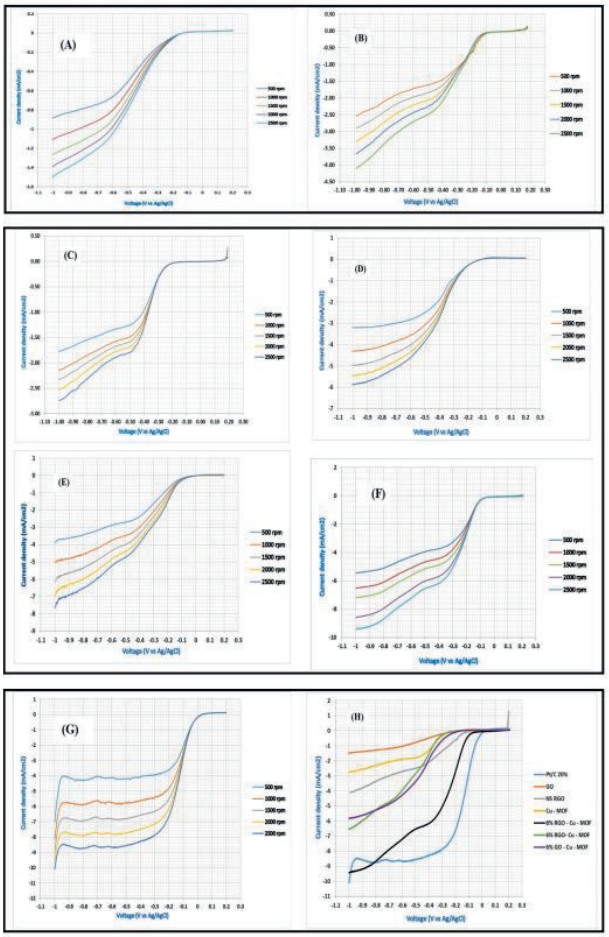


Figure 5. The LSV curves of synthesized electrocatalysts, and 20 wt% Pt/C commercial catalyst: (A) GO; (B) NS – RGO; (C) Cu – MOF; (D) 6% GO - Cu – MOF; (E) 6% NS – RGO - Cu – MOF; (F) 8% NS – RGO - Cu – MOF; (G) 20 wt% Pt/C; (H) comparison of electrocatalysts.

At the final stage of investigation, the electrochemistry activity of 8% NS – RGO - Cu – MOF is compared to the previous studies, which the types of MOFs are utilized at the cathode side (ORR). Table q shows the obtained onset potential for this study and other investigation. In this study, the best electrochemistry activity is related to 8% NS – RGO - Cu – MOF (-0.06 V vs. Ag/AgCl). It can be concluded that by using the NS – RGO as

support, which has a high electron conductivity, the performance and activity of Cu – MOF have been improved considerably.

Table 1. The obtained onset potentials for this investigation and prior works at 0.1 M KOH.

Name of	*Onset potential*	*Authors*	*Ref.*
8% NS – RGO - Cu – MOF	-0.06	-	This work
Fe-BTC	-0.12		[7]
(Fe/Co)-BTC	-0.13		[9]
$Cu_2(TMBDI)(H_2O)_2$	-0.13	Jiang, Min, et al.	[21]
Co-Oxybis (benzoic acid)	-0.197	Fan, Tianyu, et al.	[11]
Co/MIL - 101(Cr) - O	-0.12	Xiaobo He et al.	[8]
Co/MIL - 101(Cr)	-0.08	Xiaobo He et al.	[8]

IV. CONCLUSIONS

In summary, by applying the simple hydrothermal methods, several free- platinum electrocatalysts are fabricated for ORR applications. The focus of this investigation is to synthesize the carbon-based materials (GO and NS – RGO) incorporated Cu – MOF composite for ORR. Amongst the synthesized electrocatalysts, 8% NS – RGO - Cu – MOF has the best electrochemistry activity. The onset potential and electron transferred number of this catalyst are found to be -0.06 V vs. Ag/AgCl and 3.35, respectively. Actually, it tends to favor the 4e- pathway for oxygen reduction reaction. The performance of this catalyst is very close to the 20 wt% Pt/C commercial catalyst. Also, the synthesized electrocatalysts have an acceptable durability. This investigation indicated that the metal – free NS – RGO catalyst can considerably improve the poor electron conductivity of MOFs. Also, applying the heat-treating method can enhance the activity and conductivity of MOF. The electrochemistry activity of wt. % of NS – RGO – Cu – MOF electrocatalysts is promising and competitive with 20 wt% Pt/C commercial catalyst. In other words, the composite of carbon - based materials and MOFs can help to development of fuel cells technology.

REFERENCES

1. Liu, X., et al., *One-step synthesis of shell/core structural boron and nitrogen co-doped graphitic carbon/nanodiamond as efficient electrocatalyst for the oxygen reduction reaction in alkaline media.* Electrochimica Acta, 2016. **194**: p. 161-167.
2. Liu, X.-H., et al., *Well-defined metal–O6 in metal–catecholates as a novel active site for oxygen electroreduction.* ACS applied materials & interfaces, 2017. **9**(34): p. 28473-28477.
3. Li, H., et al., *One step in-situ synthesis of Co@ N, S co-doped CNTs composite with excellent HER and ORR bi-functional electrocatalytic performances.* Electrochimica Acta, 2017. **247**: p. 736-744.
4. Chung, H.T., et al., *Direct atomic-level insight into the active sites of a high-performance PGM-free ORR catalyst.* Science, 2017. **357**(6350): p. 479-484.
5. Chai, L., et al., *Bottom-up synthesis of MOF-derived hollow N-doped carbon materials for enhanced ORR performance.* Carbon, 2019. **146**: p. 248-256.
6. Gonen, S. and L. Elbaz, *Metal organic frameworks as catalysts for oxygen reduction.* Current Opinion in Electrochemistry, 2018. **9**: p. 179-188.
7. Song, G., et al., *Preparation of MOF (Fe) and its catalytic activity for oxygen reduction reaction in an alkaline electrolyte.* Chinese Journal of Catalysis, 2014. **35**(2): p. 185-195.
8. He, X., F. Yin, and G. Li, *A Co/metal–organic-framework bifunctional electrocatalyst: the effect of the surface cobalt oxidation state on oxygen evolution/reduction reactions in an alkaline electrolyte.* international journal of hydrogen energy, 2015. **40**(31): p. 9713-9722.
9. Wang, H., et al., *Preparation, characterization and bifunctional catalytic properties of MOF (Fe/Co) catalyst for oxygen reduction/evolution reactions in alkaline electrolyte.* International journal of hydrogen energy, 2014. **39**(28): p. 16179-16186.
10. Sheelam, A. and K. Ramanujam, *Metal-Organic Complexes,[Co (bpy) 3](NO3) 2 and [Co (bpy) 2NO3] NO3· 5H2O, for Oxygen Reduction Reaction.* Journal of The Electrochemical Society, 2017. **164**(9): p. F1022-F1029.
11. Fan, T., et al., *A metal–organic-framework/carbon composite with enhanced bifunctional electrocatalytic activities towards oxygen reduction/evolution reactions.* International Journal of Hydrogen Energy, 2017. **42**(27): p. 17376-17385.
12. Fang, Y., et al., *Self-assembly of cobalt-centered metal organic framework and multiwalled carbon nanotubes hybrids as a highly active and corrosion-resistant bifunctional oxygen catalyst.* Journal of Power Sources, 2016. **326**: p. 50-59.
13. Jahan, M., Z. Liu, and K.P. Loh, *A Graphene oxide and copper-centered metal organic framework composite as a tri-functional catalyst for HER, OER, and ORR.* Advanced Functional Materials, 2013. **23**(43): p. 5363-5372.
14. Song, B.-Y., et al., *Achievement of a novel porous non-noble-metal catalyst with excellent oxygen reduction reaction activity: Promoting the commercialization of alkaline fuel cells.* Journal of Cleaner Production, 2020. **249**: p. 119314.
15. Gokhale, R., et al., *Novel dual templating approach for preparation of highly active Fe-NC electrocatalyst for oxygen reduction.* Electrochimica Acta, 2017. **224**: p. 49-55.
16. Jabbar, A., et al., *Electrochemical deposition of nickel graphene composite coatings: effect of deposition temperature on its surface morphology*

and corrosion resistance. RSC advances, 2017. **7**(49): p. 31100-31109.
17. Soomro, S.A., et al., *Improved Performance of CuFe2O4/rGO Nanohybrid as an Anode Material for Lithium-ion Batteries Prepared Via Facile One-step Method.* Current Nanoscience, 2019. **15**(4): p. 420-429.
18. Liu, G., et al., *Heat-pulse assisted NH 3 gas sensing based on cuprous oxide nanoparticles anchored on reduced graphene oxide nanosheets.* Journal of Materials Science: Materials in Electronics, 2018. **29**(4): p. 3317-3325.
19. Bag, S., et al., *Nitrogen and sulfur dual-doped reduced graphene oxide: synergistic effect of dopants towards oxygen reduction reaction.* electrochimica acta, 2015. **163**: p. 16-23.
20. Gurunathan, S., et al., *Green synthesis of graphene and its cytotoxic effects in human breast cancer cells.* International journal of nanomedicine, 2013. **8**: p. 1015.
21. Jiang, M., et al., *Oxygen reduction in the nanocage of metal–organic frameworks with an electron transfer mediator.* Journal of Materials Chemistry A, 2014. **2**(15): p. 5323-5329.

Chemical Modification of Starch for Food Nanomaterial Applications

F. SAMAN[1], O. ARSLAN[1] and Y. ABALI[2]
[1] Istanbul Sabahattin Zaim University, Food Eng. Dept., 34303, Istanbul / Turkey, filiz.saman@std.izu.edu.tr , o.arslan@izu.edu.tr
[2]Manisa Celal Bayar University, Chemistry Department, 45140, Manisa / Turkey, Yüksel.abali@cbu.edu.tr

Abstract **- Starch is a common food ingredient which composed of amylose chains formed by the combination of approximately 200 glucose units linked by a linear α-1,4 bond. In addition to that starch contains amylopectin structures consisting of approximately 20 glucose units with α-1,6 branching in amylose chains. Chemical modification of the starch is realized for many different applications in nutritional purposes, surface adsorption and water regulation. The -OH groups of glucose monomers in the starch structure play an important role for the hydrogen bonding, covalent modification and hydrolysis condensation reactions. From this perspective, Si quantum dots which were synthesized by sol-gel technique, was interacted with starch molecules. During the synthesis of Si quantum dots hydrolysis and condensations reactions take place forming a surface modified Si quantum dots emitting a well defined visible light detected by photoluminescence measurements. Chemically modified starch shows visible light emission which was confirmed by photoluminescence peaks and Fourier transformed infrared spectrums. Obtained fluorescence materials can be utilized for food nanomaterial applications such as food barcods, fluorescence sensors and exceptional starch based applications.**

Keywords **- Modified starch, fluorescence, quantum particle, food nanomaterial**

I. INTRODUCTION

THE stored structure of carbohydrates formed as a result of photosynthesis in plants is called starch. Starch derivatives is utilized mainly in the food industry (raw materials or thickeners, gelling agents and encapsulation agents), paper industry (as wet end additives for dry strength, surface sizes and coating binders), as adhesives (corrugated cardboard, bags, bottle labeling, lamination, envelopes, pipe wrapping and wallpaper pastes), warp sizing of textiles and glass fiber sizing [1], [2].

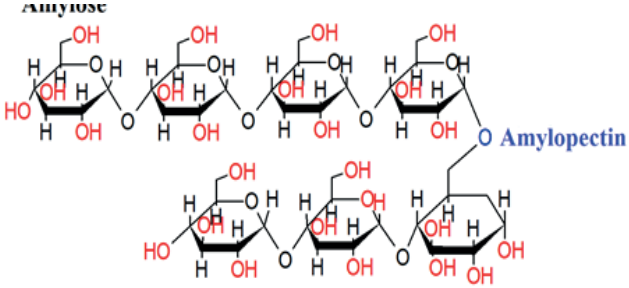


Figure 1: Chemical binding in of starch molecules

The functional properties of starches are based on the molecular structure of amylose and amylopectin in their chemical composition [3]. The combination of glucose units with linear α-1,4 bond forms the amylose structure. The glucose groups connected to this amylose structure by branching α-1,6 form the amylopectin structure. Glucose molecules contain hydroxyl groups that can react chemically [4]. By modifying the starch through hydroxyl groups of these glucose molecules, a new novel feature can be attached to the starch. Numerous studies were conducted for starch modification from past to present [5]–[7].

It is inevitable that modification techniques change with the developing technology and in this direction, the expectations from modified materials will also vary. Today, with the development of nanotechnology, materials can be modified by coating the surfaces of organic materials [8]. In this way, these organic materials can exhibit chemical and physical properties that they normally do not carry. Nano structures are structures below 100 nm. These structures show different properties according to their size and quantization situations.

Quantum particles are structures that are synthesized below 10 nm in size and exhibit special behaviors depending on their size due to quantized energy levels. These particles can be organic or inorganic structures together with surface modifications. There are basically two different methods for the synthesis of quantum particles. The first is to obtain quantum dots by shrinking amorphous matter from top to bottom. The second is the method of synthesis from molecule to material from the bottom up.

In this study, starch was aimed to realize the modification with quantum dots in order to expand the application area. Also this modification will increase the functionality of starch and its derivatives.

II. MATERIAL AND METHODS

A. Production of SiQDs

For the synthesis of Si quantum particles, various reductive ligand molecules and one aminosilane are prepared in different proportions and the mixture was exposed to UV rays. Thus, Si quantum particles were obtained. By changing the exposure time Si QD features can be varied. At this stage, the surface may show covalent bonding with other structures [9-10].

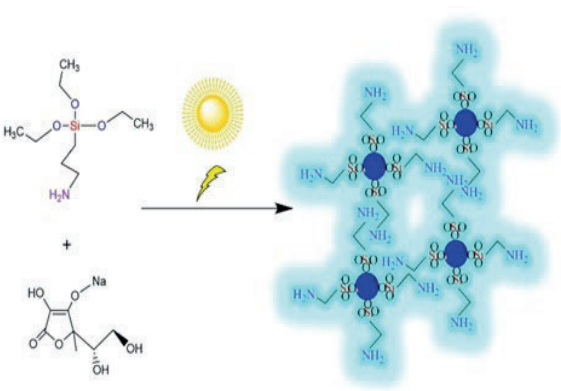


Figure 2: Production of Si QDs from the precursors

B. Modification of starch

Synthesized Si QD's carry -OH and OC2H5 groups appearing on the surface of the particles during the synthesis. These groups are vectorized for the starch modification in aqueous solution. According to the produced procedure, certain amount of starch is placed into the test tube and QD solution is added onto this dried powder. After mixing for 10 minutes, obtained modified starch molecules were dried and stored. Since hydrolysis and condensation reactions basically govern the modification, this covalent bond examination was conducted by FT-IR analysis. Figure 2 accordingly shows the surface properties and functional groups on Si QDs. Also Figure 2 shows the surface properties and anchor points after the fabrication of quantum particles.

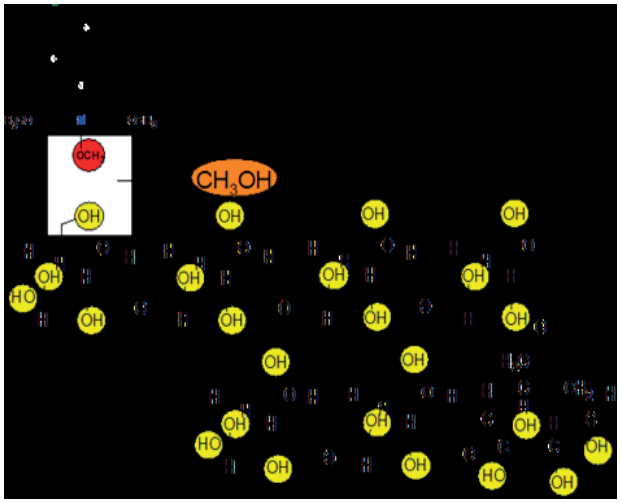


Figure 3: Interaction of quantum particles in starch modification

In addition, modification features are also detectable as shown in Figure 3. For the all anchor –OH positions, possible condensation reactions were shown. Thus, –OH positions are determined as an ideal modification point for Si QDs. After the side products of condensation is removed (Figure 4) volatile EtOH/MeOH for different reaction systems, total reaction is completed.

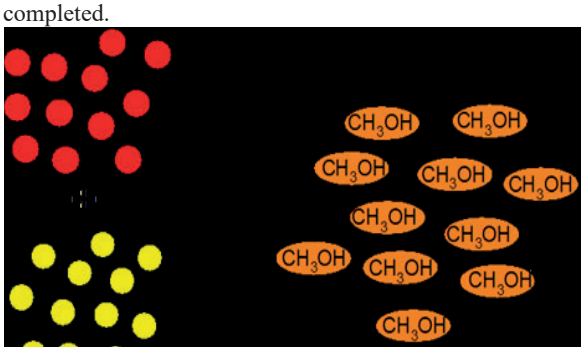


Figure 4: Demonstration of the scheme of binding of quantum particles to -OH regions in starch modification (yellow: -OH positions, red: -OCH3 groups)

C. Applications of modified strach

In order to examine the food availability of the modified starch, a guda product was prepared and a formulation for water pudding was completed. The fluorescence properties of the obtained water pudding were examined.

III. RESULTS

For the synthesis of quantum particles, a certain amount of starting silane material was taken and solvent and reducing agent were added on the basis of volume ratios. First of all, the amount of reactant material is optimized and it is aimed to obtain a sustainable quantum dot material. Additionally, interaction period with the UV lamp was optimized. For this purpose, the same ratios were interacted with different amounts of UV light and their fluorescence properties were examined.

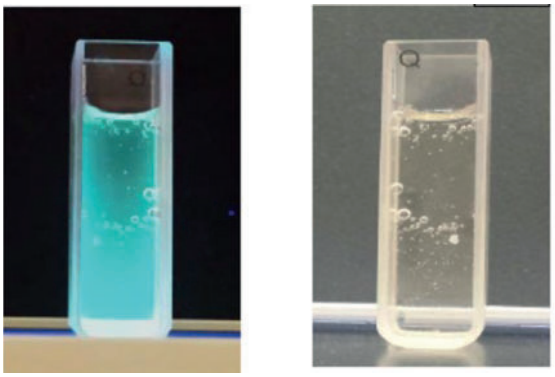


Figure 5: Si Quantum Dots under excitation (left) under daylight (right)

Technically Si QD interaction with the starch is conducted in a test tube. But results showed that obtained products may inherit some weight losses during the procedure and therefore plastic tubes or teflon mediums seem more suitable for the

reaction. Modification process was performed with different amounts of Si QDs to maintain a homogenous, detectable fluorescence intensity. SEM investigation was revealed that prior to modification standart hollow structures of starch is easily detected. One hollow structure exhibits a stable single item having the smooth surface properties. Roughly these hollow starch structures are 15-20 micrometer in size. At first look, starch hollow structures are look like micropotatoes. As mentioned surface roughness is extremely low which may be due to the extreme polar character of the monomolecular structure of glucose. Since –OH groups area extreme, these polar structure forces the starch molecules to form a barrier to each other establishing a closed hollow structure decreasing the thermodynamic instability.

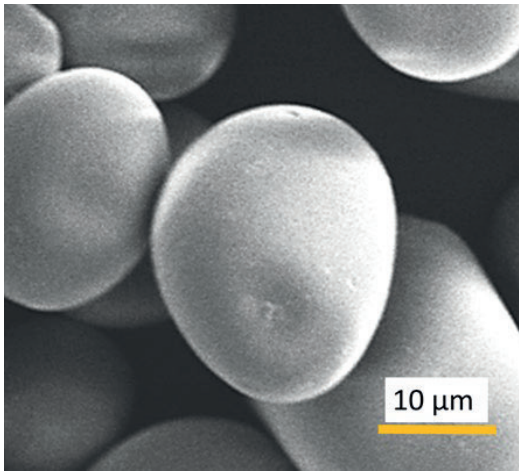


Figure 6: SEM investigation of the raw starch hollow structures

During the surface modification of these starch hollow structures, compact appearance and surface features vary significantly. SEM investigation shows that potatoes like shapes of starch demolishes and a random shape behaviour is observed.

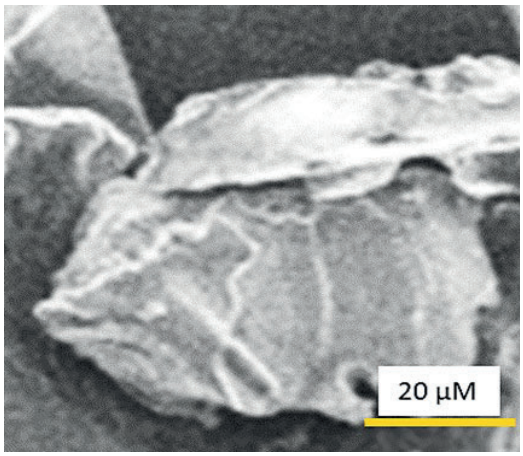


Figure 7: SEM image of modified starch

This random shaped new modified structure of starch reveals that modification of the raw starch was resulted with the formation of random geometries representing the succesfull modification. Smooth surfaces of the starch was varied and structures which are greater than raw starch was obtained. Closed observation unveils that modified structure shows plate like morphology and sizes are randomly increased. Different and irregular sizes inform us that modification effects more than one hollow structure of the raw starch. Consequently, one hollow structure interacts with more than one Si QD and this causes attaching raw starch geometries to each other resulting bigger, tougher organic clusters with fluorescence features.

Once starch structure are modified, physical appearance of the starch does not vary since Si QD concentration on the starch structures are extremely low but since fluorescence is a very sensitive method it still provides the fluorescence feature which may be utilized for different applications. Figure 8 shows the physical appearance of the modified starch structure under daylight and excitation.

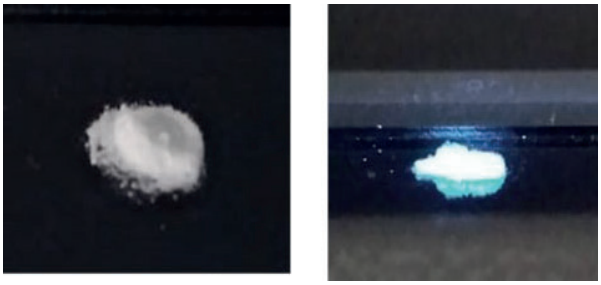


Figure 8: The image of the modified starch in daylight and under excitation

During the fabrication procedure of the Si QD's varying periods of UV illumination was applied causing small shiftings in the PL peaks. Therefore illumination was extended till 60-90 minutes to observe the variations. FT-IR investigation for starting precursor AMEO and potatoe starch was presented in Figure 9.

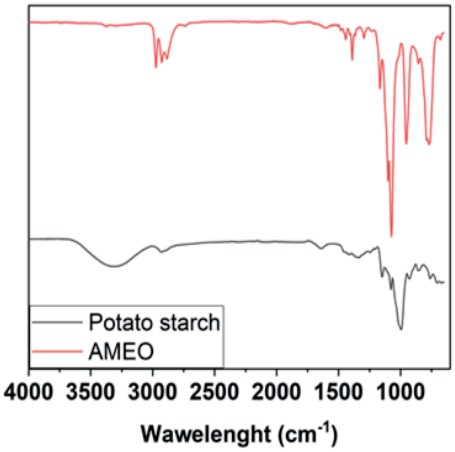


Figure 9: FT-IR analysis of starch and beginning precursor

-OH groups of the starch is clearly seen at 3300 cm-1 and O-

C peak appears at around 1000 cm-1. In Si precursor N-H peaks appear 3300-3200 cm-1 region. Also Si-O and C-O peaks are visible at around 1500-1000 cm-1. This values clearly confim that starch and Si precursor are readily reactable for the formation of visible light emitting starch structures. Especially NH2 functional groups stay on the obtained Si QD's surface providing a special version of the QD's. This can provide selective sensor applications such as Meisenheimer complex formation during the interactions.

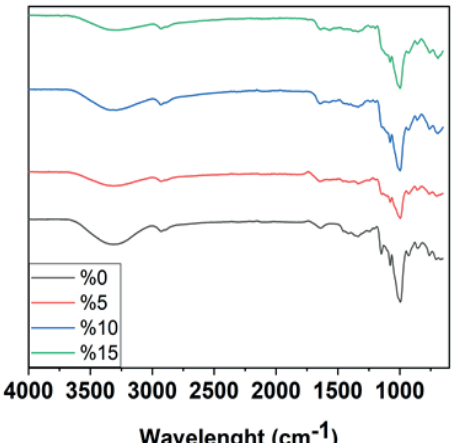


Figure 10: Different loadings of Si QD's on starch

Figure 10 reveals the FTIR investigation of different amount of Si QD loadings onto the starch structures. As the amount of the Si QD loading is increased, –OH functional group peaks intensity is accordingly decreased meaning multiple interaction of the Si QD with starch surface. Additionally, C-O peaks at around 1000 cm-1 changes relatively. Since NH2 groups are present on the starch due to the new loaded Si QDs N-H peak with C-H peaks seen after 3000 cm-1.

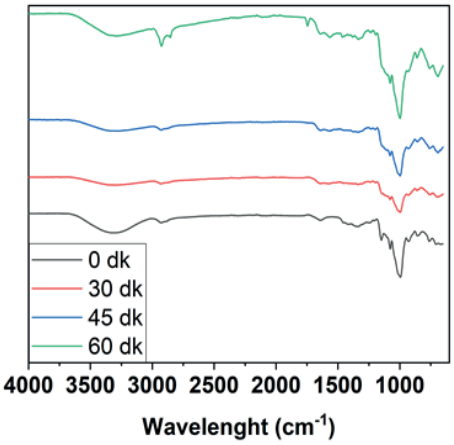


Figure 11: FT-IR investigation with different sizes of Si QD's

As mentioned earlier, varying periods of UV light was applied during the synthesis of Si QD's. When these QD's were interacted with starch no huge difference was detected. Interestingly –OH groups are still clearly visible but NH strecthings do not increase on the starch structure.

Obtained visible light emitting starch structure was utilized for the food products. This modified starch became an ingredient in water pudding. Obtained water pudding was investigated as fluorescent food. Also excited pictures of water pudding clearly shows that fluorescence features of the pudding is present which may be utilized for varied examples.

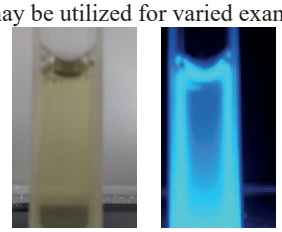


Figure 11: The image of the water pudding in daylight and excitation

As a result, Si QDS were produced with facile and easy process to modifiy the starch molecules. Chemically modified starch emits visible green light which can be utilized for varying applications. Physical and chemical investigation clearly showed that modified starch connects to the Si QD's which provide a stable composition.

ACKNOWLEDGMENT

We gratefully acknowledge Manisa Celal Bayar University-DEFAM facilities and Istanbul Sabahattin Zaim University-Halal Food Laboratory This project was supported by Istanbul Sabahattin Zaim University, IZU-PRO Projects (Project Number: İZÜPRO-2020-09).

REFERENCES

[1] J N BeMiller ve R L Whistler, Starch: Chemistry and Technology Academic Press, 2009
[2] M W Rutenberg ve D Solarek, "Starch Derivatives: Production and uses", Starch: Chemistry and Technology, Elsevier, 1984, page 311-388
[3] Georgiana Gabriela Codină, Dumitru Zaharia, Sorina Ropciuc, Adriana Dabija, Influence of magnesium gluconate salt addition on mixing, pasting and fermentation properties of dough, The EuroBiotech Journal, Vol 1, Issue 3, 2017
[4] R H Moffett, "Modified starch and process therefor", US6699363B2, Mar 02, 2004
[5] C P Patel ve R E Pyle, "Cationic starch", US3417078A, Ara 17, 1968
[6] H Zhang, R Wang, Z Chen, ve Q Zhong, Enzymatically modified starch with low digestibility produced from amylopectin by sequential amylosucrase and pullulanase treatments, Food Hydrocoll , 95, 195-202, 2019
[7] Olivia V López, Noemí E Zaritzky,María A García, Physicochemical characterization of chemically modified corn starches related to rheological behavior, retrogradation and film forming capacity, Journal of Food Engineering, Volume 100, Issue 1, September 2010, Pages 160-168
[8] H Erdem Çamurlu, Sanjay Mathur, Osman Arslan, Esin Akarsu, Ceram Int 42, 6312-6318, 2016
[9] O Arslan, Z Aytac, ve T Uyar, "Fluorescent Si QD decoration onto a flexible polymeric electrospun nanofibrous mat for the colorimetric sensing of TNT", J Mater Chem C,5, 1816-1825, 2017
[10] O Arslan and Y Abali, "Molecular Control and Surface Modification and Inorganic Organic Hybrid, Functional Nanomaterials", Soma Mesl High School Tech Science Journal, c 2, pp 26, Art sy 26, Nov 2018

Carbon Derivative Supported Triple Metal Electro Sensor for Citric Acid Detection in Food Samples

Müge Yayla[1*], Hilal Çelik Kazıcı[2], Hilal Demir Kivrak[2]

[*1] *Department of Chemistry and Chemical Processing Technologies, Muradiye Vocational School, Van Yüzüncü Yıl University, 65080 Van, Turkey*
mugeyayla@yyu.edu.tr
[2] *Department of Chemical Engineering, Faculty of Engineering, Van Yüzüncü Yıl University, 65080 Van, Turkey*
hilalkazici@yyu.edu.tr
[2] *Department of Chemical Engineering, Faculty of Engineering, Van Yüzüncü Yıl University, 65080 Van, Turkey*
hilalkivrak@yyu.edu.tr

ABSTRACT

A PdRuCr-based sensor that detects the citric acid (CA) preservative in different foods has been developed. Carbon nanotube (CNT) based PdRuCr catalysts were synthesized by sodium boron hydride ($NaBH_4$) method. The activities of the electrocatalysts were carried out using electrochemical measurements such as cyclic voltammetry (CV) and differential pulse voltammetry (DPV) methods. It was studied by adding Cr in different proportions to the structure for PdRuCr / CNT, which is a three metal catalyst.. The concentration study was carried out between 0 and 10 mM. The optimum concentration for citric acid was determined as 5 mM. Scan rate was carried out to determine the reaction mechanism. Analytical characterization of the electrocatalyst was determined using the differential pulse voltammetry technique for the specification limit and sensitivity values of each preservative. The lowest limit ranges (LOD) and senstivity were determined for citric acid (CA), the specification limit values of the sensors in ppm were calculated as 2.16 mgL-1 and 3392,958 µA/mM.cm^2, respectively. Analyzes were performed on different foods for CA preservative. These different foods: grape vinegar, grape vinegar and orange soda.

Key words: citric acid, electro-sensor, food sensor

I. INTRODUCTION

Assuring food quality and safety has become a subject of major concern for authorities and professionals of the food supply chain. Food quality relates primarily to characteristics such as appear-ance, texture, taste, smell, nutritional value content, which have to be ensured to guarantee both nutritional quality and acceptability of the delivered foods by the consumers. Analysis of food composi-tion helps controlling that all the desired constituents, including natural components (eg, sugars, amino acids, and alcohols) and ad-ditives (eg, vitamins and minerals), are present in a suitable range of concentration. Food analysis is also of uppermost importance to protect the consumers against any adulteration, spoilage, or con-tamination [1].

One of the key challenges to both control foodstuff composition and detect chemical or biological contaminants is the availability of selective, sensitive, and reliable analytical methods. To date, most analyses are performed by skilled personnel in accredited laboratories using sophisticated and costly instrumentation such as liquid or gas chromatography coupled with mass spectrometry. Among the emerging and innovative tech-nologies proposed for food analysis, biosensors are presented to date as very promising low-cost screening tools usable as comple-mentary devices to conventional methods. Biosensors could help to select a rather limited number of suspect samples that would be further analyzed for confirmation by conventional techniques, reducing the cost and time of analysis [2].

In this study, three metal catalysts were synthesized and electrochemical studies were carried out on citric acid, a food preservative. The surface morphology of the obtained nanomaterials were investigated by cyclic voltammetry (CV) technique. As a result of the studies, the most sensitive, reliable, highly repeatable and economical sensor was obtained.

II. EXPERIMENTAL METHODS

In this study, PdRuCr / CNT nano catalysts were synthesized by sodium boron hydride reduction method.

III. RESULTS AND DISCUSSION

The buffering effect of the three metal catalyst was investigated. According to the data obtained as a result of the experiments, the best buffer value for 7% of the three metal catalyst was determined to be 0.1 M H2SO4 buffer.

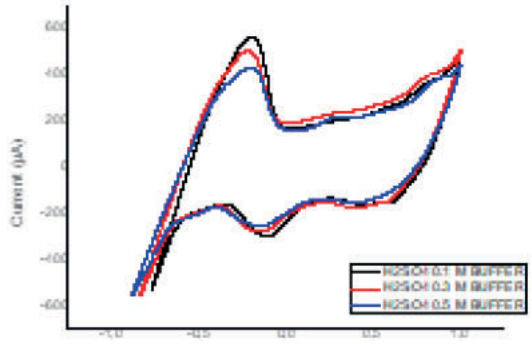


Figure 1: Buffering effect of 7% PdRuCr / CNT catalyst for 5 mM citric acid.

Concentration study was carried out on the catalyst determined for citric acid. The results of this study are shown in the figure 3.

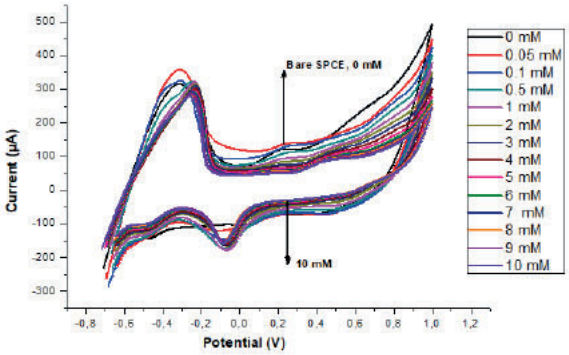


Figure 2: The cyclic voltammogram of 7% PdRuCr / CNT catalyst taken in 0.1 M H2SO4 at a concentration range of 0 - 10 mM of citric acid and at 50 mVs-1.

Another important advantage of the alternating voltammetry technique is that the method can be applied at different scanning rates. Thus, issues related to the stability of the intermediate products formed by the electrode reaction can be determined. Besides the intermediate products, it is possible to determine the chemical reaction events that affect the adsorption, diffusion and electron transfer reaction.

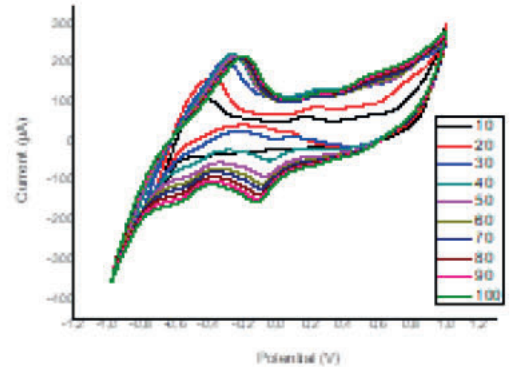


Figure 3: Effect of scanning rate on oxidation of citric acid in PdRuCr / CNT (10-20-30-40-50-60-70-80-90-100 mVs-1).

Oxidation peak currents in CVs taken at different scan rates between -1 V and +1 V potential at 10 - 100 mVs-1 in 0.1 M H2SO4 buffer medium were plotted against the square root of the scan rate. As can be seen from the line equations, the correlation numbers close to 1 indicates that the currents are diffusion controlled.

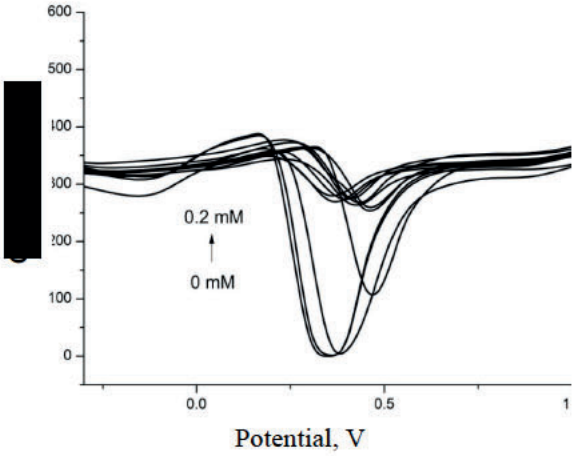


Figure 4: Differential pulse voltammetry (DPV) in a 0.1 M H2SO4 buffer solution in citric acid concentrations ranging from (12.5 µM to 200 µM).

Analytical characterization for citric acid was carried out by differential pulse voltammetry (DPV). With this voltammetry, a calibration graph is given to obtain data about the detection limit and sensitivity of the sensor (Figure 7).

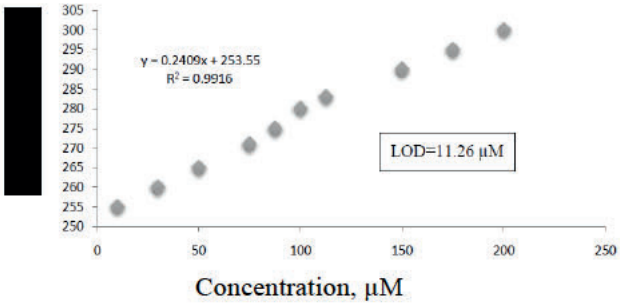


Figure 5: Citric acid calibration graph.

While determining the limit of determination (LOD) with calibration graphics drawn from DPV voltammograms, regression analysis was performed and it was observed that the concentration and current densities were linear and the number of correlations was high.

The sensors' specification limit value in ppm has been calculated to be 2.16 mgL-1. The sensitivity of the obtained sensor was calculated as 3392,958 µA / mM.cm2.

The parameters (LOD, sensitivity, correlation number) calculated from the created calibration graph proved that

the electroanalytical method has high accuracy and is applicable.

IV. REFERENCES

[1] Dridi, F , Marrakchi, M , Gargouri, M , Saulnier, J , Renault, N J , Lagarde, E , 2017 Nanomaterıal-based electrochemıcal bıosensors for food safety and qualıty assessment Nanomaterıal-based electrochemıcal bıosensors Chapter 5, 167-204
[2] McGrath, T F , Elliott, C T , Fodey, T L , 2012 Biosensors for the analysis of microbiological and chemical contaminants in food Anal Bioanal Chem 403, 75–92

Engineered Fabrication of the SiO_2 Nanoparticles by Tailored Stoeber Process

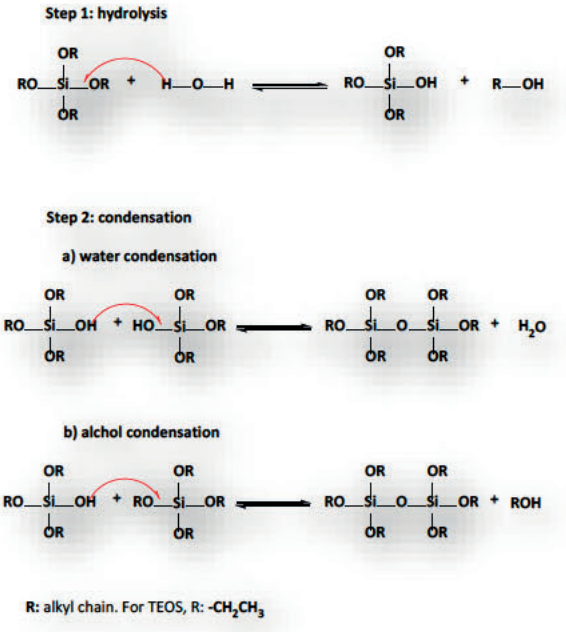
Bilge Boylu[1], Osman Arslan[1], Yüksel Abalı[2]
[1] Istanbul Sabahattin Zaim University, Food Eng. Dept., 34303, Istanbul / Turkey,
o.arslan@izu.edu.tr, boylu.bilge@std.izu.edu.tr
[2] Manisa Celal Bayar University, Chemistry Department, 45140, Manisa / Turkey,
yuksel.abali@cbu.edu.tr

Abstract **- As one of the most prominent SiO_2 nanoparticle synthesis method, Stoeber technique was modified and investigated for a better understanding of the nanoparticle formation and surface analysis. For this synthetic method, tetraethoxysilane presursor was utilized by minor modifications in the fabrication route. Especially concentration effect was studied thoroughly. By keeping the pH value and other parameters same, different concentrations of the tetraethoxysilane precursor provided diverged results about size and surface characteristics. As sol-gel reaction mechanism implies, basic catalysis and temperature plays an important role for the synthesis of SiO_2 nanoparticles. This modified methodology also allows us to imply a surface and particle size control for the various applications of SiO_2 nanoparticles. Size distribution and EDX analysis unveiled that a regular size control is possible with high sensitivity and there is no other impurities or atomic entities in the obtained nanoparticles. According to the Fourier Transformed Infrared spectroscopy, surface contains –OH groups and particles carry water molecules which is a sign for SiO_2 particles that they can be effective for surface humidity adsorption in different applications. Monodispersed, spherically shaped and aglomeration free/low agglomerated particles were obtained and statistically investigated for possible applications such as sensors, theranostics, humidity adsorption and chemical delivery.**

Keywords **- Nanoparticle, Stoeber process, sol-gel, SiO_2, surface modification**

I. Introduction

SiO_2 became a center nanomaterial for the researchers in many different applications recently. As a nanofiller, SiO_2 embedded polymer nano-composites were reported in the last few years with significant increase [4]. Advantageous mechanical and thermal properties of composite materials with SiO_2 filled nanoparticles in polymer matrix were reported [3,7]. In another study, it was found that modified SiO_2 nanoparticle-based films shows antimicrobial activity [5]. Additionally, fabrication of superhydrophobic surfaces with SiO_2 nanoparticle containing thin films and their potential applications (anti-corrosion, self-cleaning, anti-icing, anticontamination, non-stick surfaces, etc.) were multiplied enormously [9]. In a sample study, we see that self-cleaning, water-resistant hybrid nanofiber mats are obtained with modified cellulose acetate nanofibers that separate oil-water structures from each other [1].



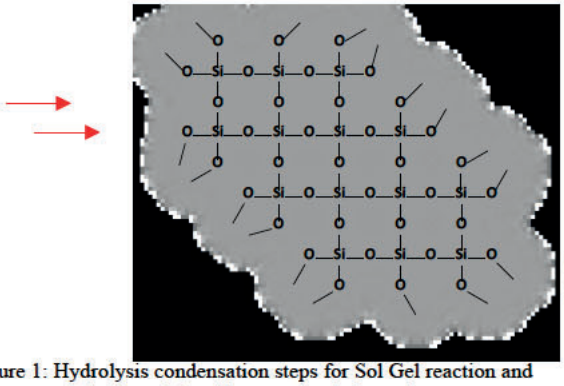


Figure 1: Hydrolysis condensation steps for Sol Gel reaction and image of the silicate network formed

Since the physical, chemical and optical properties of silica particles depend on their size, surface features and monodispersity, effective methods of preparing particles of different sizes [6] with controlled chemical character requires a particular attention. The size and monodispersity of silica particles has a major impact on the quality of these structure.

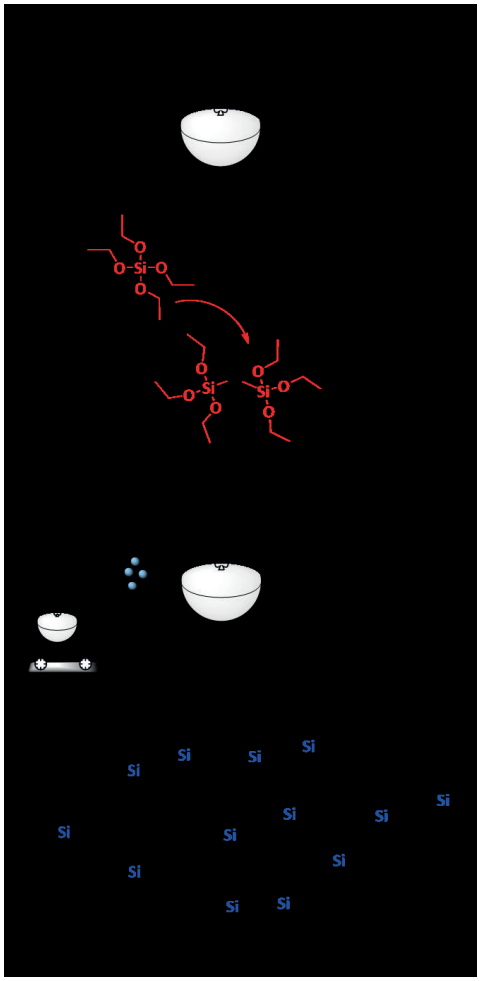


Figure 2: Stoeber process for SiO_2 synthesis

Stoeber et.al. produced a synthesis route to produce silica nanoparticles with variable particle sizes between 5-2000 nm using the sol-gel method in alkaline medium [6]. This technique is considered to be a simple and effective way to fabricate uniform silica spheres since the reaction conditions can be controlled with details [10]. Generally, silicon alkoxide (usually tetraethyl orthosilicate -TEOS) is hydrolyzed in the presence of an alcohol under the basic catalyst such as ammonia [8]. Sol-gel process is mainly based on two reactions, hydrolysis and condensation of precursors of metal alkoxides or metal salts [2]. Especially Si-O-Si bonds formed by the hydrolysis and condensation of alkoxysilanes for fabricating SiO_2. As shown in Figure 1, the development of a three-dimensional cross-linked inorganic network structure with $Si(OR)_4$ precursor is easily detectable.

In this study different concentration values were applied to the modified Stoeber method to observe the concentration effect with usual synthesis conditions. Investigation showed that size, surface character and monodispersity features can be modified and modulated for different applications.

II. MATERIAL AND METHODS

Chemicals: Tetraethoxysilane (98%) (TEOS) obtained from Acros Organics. Octyl triethoxysilane (OCTEO) were obtained by Degussa-Dynasilane. Distilled water were added for the hydrolysis-condensation reactions. Ethanol (technical), Isopropyl alcohol (99,95%), NH_3 (26%) have been received from Sigma-Aldrich. Fabricated particles were cleaned with EtOH and acetone and dried under room condiitons.

Methods and Characterization: Shimadzu-IRTracer-100 with ATR module was employed to measure the surface character of the obtained nanoparticles. Shape and monodispersity investigation were obtained by Scanning Electron Microscopy (SEM) with ZEISS GEMINI 500 device. SEM Samples were placed on Cu tape and sputtered with 3-5 nm Au prior to the analysis. Particle size analysis were investigated by counting the 50 particle and plotting their frequencies versus particles sizes. Non linear analysis were applied to find the standart deviation.

Preparations of Silica nanoparticles: Nanosized SiO_2 particles were obtained by modified Stoeber method. For the synthesis, $IPA/H_2O/NH_3$ mixture was prepared between 0-5 °C (cold trap) for a certain period of time. In a 250 mL flask, proportions of the $IPA/H_2O/NH_3$ reactants were 4/1/0.75 and this mixture were mixed at 400 rpm for 15 minutes. Then, according to the same proportion 0.35/0.70/1.05 amounts for TEOS were added in separate experiments. After the TEOS, reaction were allowed to carry on at 400 rpm for another 6 hours. At the end, solution was centrifuged (HITACHI CR22N, Japan) (10000 rpm, 5 minutes) and washed with EtOH and acetone respectively and dried at room conditions.

III. RESULT AND ANALYSIS

SiO_2 nanoparticle synthesis were conducted with sol-gel method using Stoeber technique in one pot reaction. The synthesis was carried out by a modified route with varying sialne concentrations. For the fabrication, an ambient environment were provided during the synthesis by modulating the necessary parameters. Especially at the early stages of the nucleation by hydrolysis-condensation reactions, cold trap was applied and this is the most significant parameter for the synthesis. Prepared cold trap allows us to control the propagation of the polymerization reactions and also modulate the particles sizes.

In order to analyze the effect of silane concentration, we have utilized varied amounts of silane precursor for the nanoparticle synthesis. Variation was repeated 3 times and SiO_2-1, SiO_2-2 and SiO_2-3 were obtained respectively. Obtained particles were

subjected to SEM, size definition, EDX and FT-IR analysis to investigate their chemical and physical features.

SEM image of SiO_2-1 nanoparticle was presented in Figure3. According to the image, SiO_2 nanoparticles show varying morphologies but very close to the spherical structure. Sizes of the nanoparticles do not clearly show a monodispersed behaviour but size distribution claims that the particle sizes generally focus on the 500-600 nm range (Figure 4) gap.

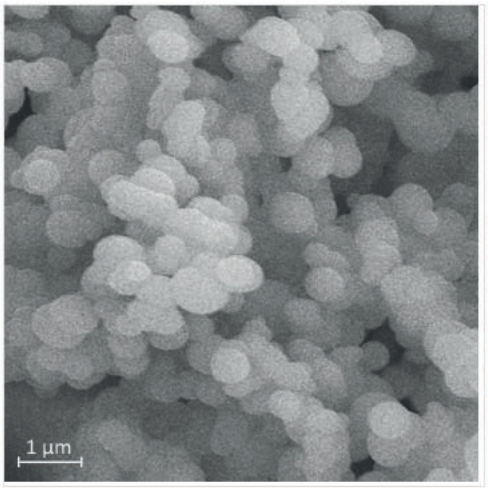


Figure 3: SEM investigation of the SiO_2 -1 EDX investigation (Figure 5) clearly shows that SiO_2 nanoparticles do not have any other cation or anions. Additionally, there is no unexpected atomic entities in the particles.

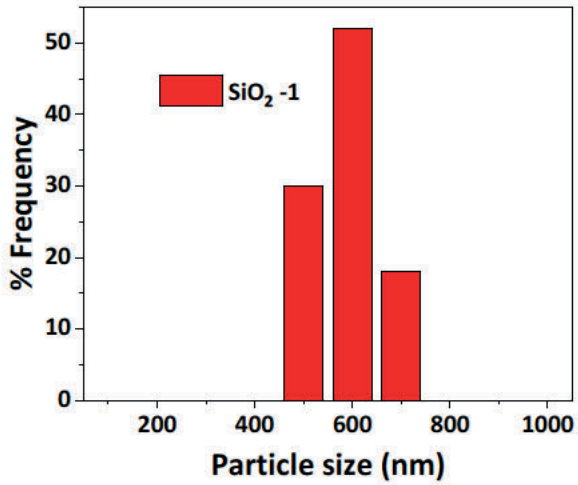


Figure 4: Size distribution of the SiO_2-1

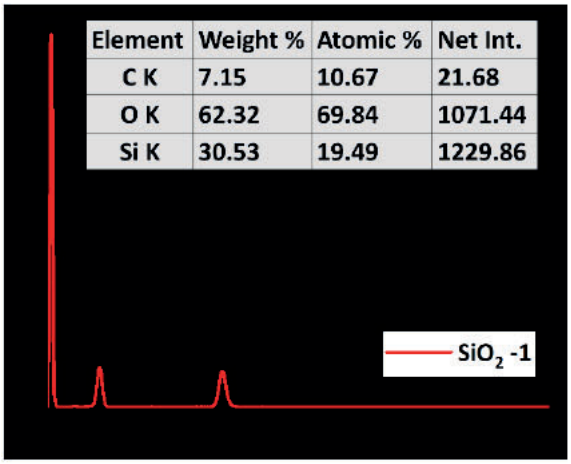


Figure 5: EDX investigation of the SiO_2-1

Similarly, when SEM image (Figure 6) of the SiO_2-2 is investigated relatively better spherical morphology was observed. Therefore, one can conclude that morphological efficiency would be better for SiO_2-2 nanoparticles. Monodispersity of the same nanoparticles seems improved due to the increased concentration. Surface character of the nanoparticles implies lower agglomeration but larger investigation may be needed. Size distribution unveils that average size is about 700-800 nm range (Figure 7). This results show 2 significant reflections. First one is, with the increasing concentration of the silane precursor, average size increases about 100 nm. Second derivation is, if there is no variation in the synthesis procedure, cold trap technique provides a straightforward and general particle size enhancement method for SiO_2 nanoparticles. EDX investigation (Figure 8) shows Si, O and C (possibly artifact) for SiO_2-2.

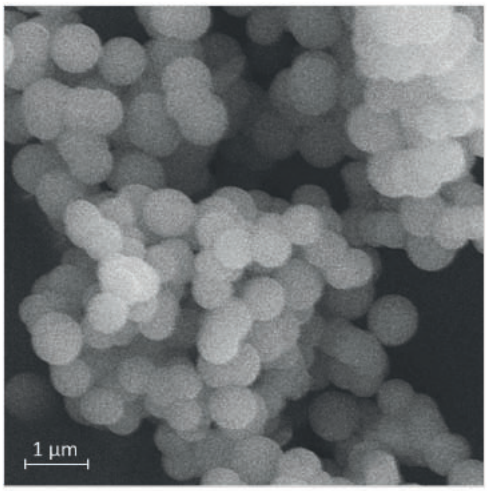


Figure 6: SEM investigation of the SiO_2-2

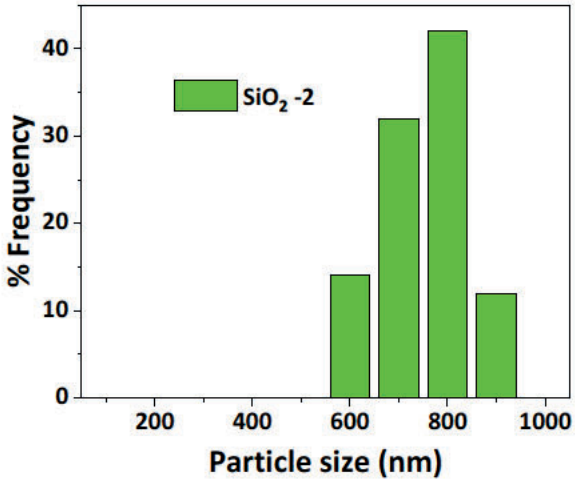


Figure 7: Size distribution of the SiO_2-2

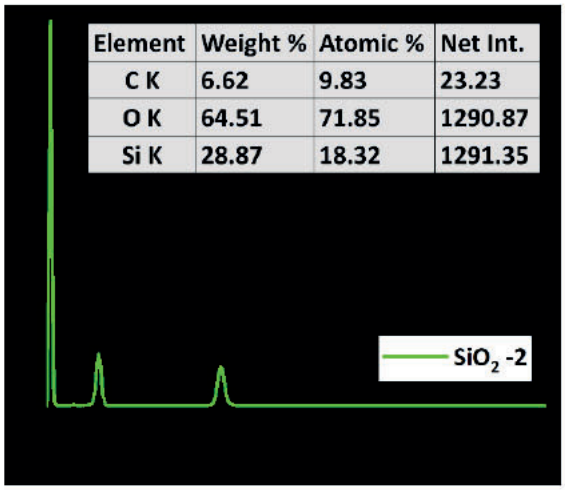


Figure 8: EDX investigation of the SiO_2-2 SEM investigation of the SiO_2-3 (Figure 9) shows similar features as in SiO_2-1. Particles are mostly spherical but monodispersity is relatively poor.

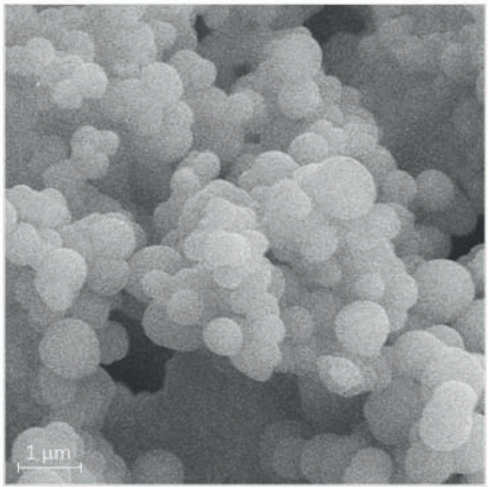


Figure 9: SEM investigation of the SiO_2-3

Particle size distribution shows that average size show 2 different peak position. First group is around 800-900 band and second is around 600 band. Frequency % is relatively higher in 800-900 band. So as mentioned earlier, enhancencement of concentration again enhances the highest frequency particle sizes. But since hydrolyses and condensation reactions occur in decreased rates, particle size enhancement gets slower. Therefore, possibly size will finalize at the 900 nm band but this will take more time than the previous examples.

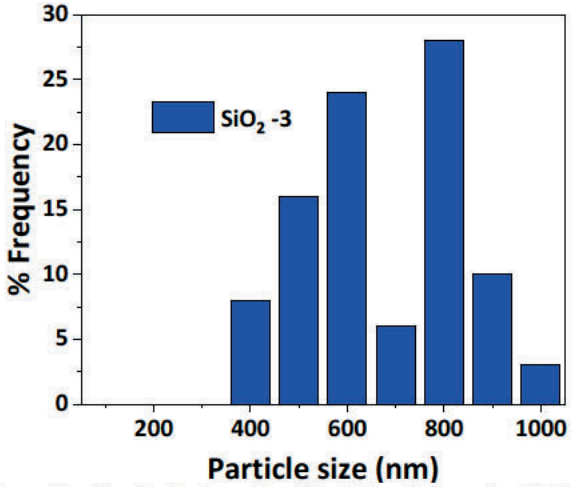


Figure 10 : Size distribution of the SiO_2-2. In addition to that EDX investigation (Figure 11) similarly shows no unexpected atom entity in the composition.

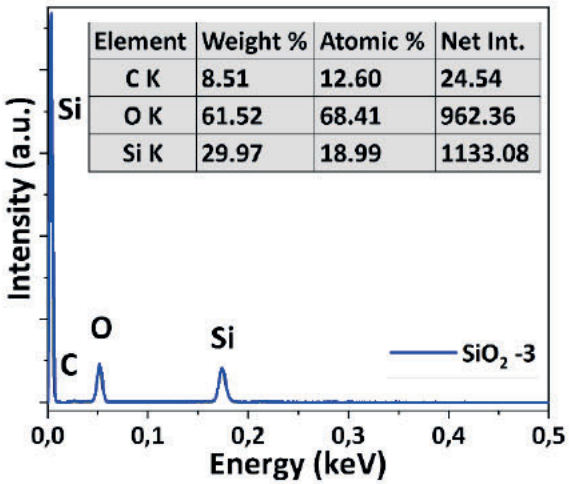


Figure 11: EDX investigation of the SiO_2-2

Surface features of the all fabricated SiO2 nanoparticles were examined by FT-IR ATR method (Figure 12). Expectedly surface contains -OH groups which can be confirmed in 3300cm^{-1} region. In addition to that Si-O band positioned at 1100 cm^{-1}. All the nanoparticle FT-IR peaks are similar which

means that all the particles urfaces have the same composition on top of the nanoparticles. Intensity of the Si-O peak at SiO_2-1 is higher than the other which claims the more Si-O band formation. As we predicted nanoaprticles grow faster in SiO_2-1 conditions and therefore we observe more Si-O peak than the others.

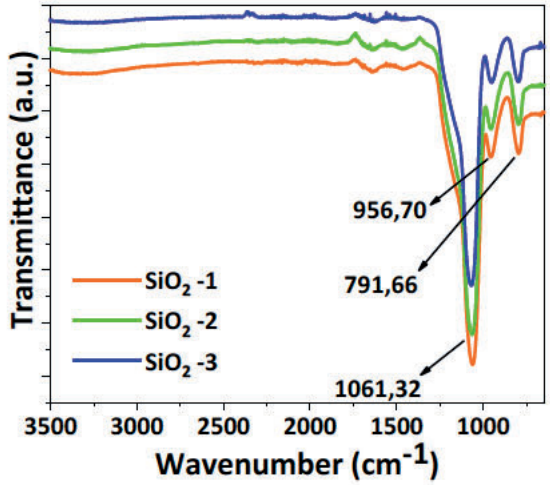


Figure 12: FT-IR investigation of the all 3 SiO_2 nanoparticles

IV. CONCLUCIONS

As a conclusion investigations and experiemental results showed that cold trap method provides a sensitive nanoaprticles size enhancement method. By a controlled concentration increase, nanoaprticles sizes are enlarged about 100 nm every time. Just in the third attempt since concentration is extremely intense, hydrolysis and condensation reactions became slower and expected higher nanoparticle size distribution frequency can not be obtained. Possibly desired nanoparticle size would be reached with elongated reaction periods.

ACKNOWLEDGMENT

We gratefully acknowledge Manisa Celal Bayar University-DEFAM facilities and Istanbul Sabahattin Zaim University-Halal Food Laboratory.

REFERENCES

[1] Arslan, O , Aytac, Z , & Uyar, T (2016) Superhydrophobic, Hybrid, Electrospun Cellulose Acetate Nanofibrous Mats for Oil/Water Separation by Tailored Surface Modification ACS Applied Materials & Interfaces, 8(30), 19747-19754 doi: 10 1021/acsami 6b05429

[2] Brinker, C J , & Scherer, G W (2014) Sol-Gel Science: The Physics and Chemistry of Sol-Gel Processing Saint Louis: Elsevier Science

[3] Chrissafis, K , Paraskevopoulos, K M , Pavlidou, E , & Bikiaris, D (2009) Thermal degradation mechanism of HDPE nanocomposites containing fumed silica nanoparticles Thermochimica Acta, 485(1), 65-71 doi: 10 1016/j tca 2008 12 011

[4] Çamurlu, H E , Akarsu, E , Arslan, O , & Mathur, S (2016) Nanocomposite glass coatings containing hexagonal boron nitride nanoparticles Ceramics International, 42(7), 8856-8862 doi: 10 1016/j ceramint 2016 02 133

[5] Hannon, J C , Kerry, J , Cruz-Romero, M , Morris, M , & Cummins, E (2015) Advances and challenges for the use of engineered nanoparticles in food contact materials Trends in Food Science & Technology, 43(1), 43-62 doi: 10 1016/j tifs 2015 01 008

[6] Meier, M , Ungerer, J , Klinge, M , & Nirschl, H (2018) Synthesis of nanometric silica particles via a modified Stöber synthesis route Colloids and Surfaces A: Physicochemical and Engineering Aspects, 538, 559-564 doi: 10 1016/j colsurfa 2017 11 047

[7] Sargsyan, A , Tonoyan, A , Davtyan, S , & Schick, C (2007) The amount of immobilized polymer in PMMA SiO2 nanocomposites determined from calorimetric data European Polymer Journal, 43(8), 3113-3127 doi: 10 1016/j eurpolymj 2007 05 011

[8] Stöber, W , Fink, A , & Bohn, E (1968) Controlled growth of monodisperse silica spheres in the micron size range Journal of Colloid and Interface Science, 26(1), 62-69 doi: 10 1016/0021-9797(68)90272-5

[9] Vidal, K , Gómez, E , Goitandia, A M , Angulo-Ibáñez, A , & Aranzabe, E (2019) The Synthesis of a Superhydrophobic and Thermal Stable Silica Coating via Sol-Gel Process Coatings, 9(10), 627 doi: 10 3390/coatings9100627

[10] Wang, X -D , Shen, Z -X , Sang, T , Cheng, X -B , Li, M -F , Chen, L -Y , & Wang, Z -S (2010) Preparation of spherical silica particles by Stöber process with high concentration of tetra-ethyl-orthosilicate Journal of Colloid and Interface Science, 341(1), 23-29 doi: 10 1016/j jcis 2009 09 018

EFFECTS OF SMELTING, RAFINATION AND HIGH PRESSURE DIE CASTING PARAMETERS ON POROSITY FORMATION IN ALUMINIUM CASTING PARTS

Leyla Şimşek[1,2], Muhammad Farooq[3], Ahmet Turan[4]

[1]Kırpart A.Ş., Gedelek Mahallesi, Gedelek Sokak, No:470, 16800, Orhangazi, Bursa, Turkey

[2]Institute of Science, Yalova University, 77200, Yalova, Turkey

[3]Department of Mechanical Engineering, University of Engineering & Technology Lahore (New Campus), Pakistan, 54890.

[4]Chemical Engineering Department, Faculty of Engineering, Yalova University, 77200, Yalova, Turkey

Abstract - Aluminium alloys have high usage rate, especially in the aviation, defence and automotive industries, where weight reduction is extremely important. Pressure die casting technology is a widely used method of casting aluminium alloys in terms of its suitability for mass production. Aluminium-silicon alloy parts, which are produced by the high pressure die casting method (HPDC), generally contain porosity in different proportions and sizes due to the discharge problems of gases trapped during the rapid injection of molten metal into the mold cavity. Because of increase in expectations and customer specifications in the casting industry, it is desired to reduce gas porosity levels to certain critical levels. Proper mold design and optimization of casting process parameters play an important role in reducing gas porosity by facilitating the evacuation of air from the mold cavity. In this study, the change of porosity formation, during the smelting, rafination and HPDC operations of AlSi9Cu3, AlSi12Cu1Fe and AlSi12Fe alloys, were investigated. Stereo microscopy and computed tomography (CT) were mainly used to characterize obtained samples.

Keywords – aluminium, computed tomography, porosity, stereo microscopy

I. INTRODUCTION

The importance of aluminium increases day by day with increase in weight reduction particularly in aviation and transportation industries [1]. The density of aluminium is around 2.7 g.cm^{-3} and it makes the Al one of the lightest structural metals after magnesium along with its suitable mechanical and physical properties like yield strength and corrosion resistance [2]. One more remarkable property of the Al is that it can easily be recycled and, its recycle results in 95% energy saving in comparison to its primary production [3].

The main reason of gas porosities and porosity-related defects in aluminium casting is the hydrogen gas which is trapped in liquid metal during casting. An increase in the amount of the hydrogen in liquid metal (as a result of the reduction of the hydrogen in water like moisture in the atmosphere) increases the ratio of porosities in casting products. Therefore, dissolved hydrogen in the liquid metal should be avoided, otherwise the mechanical properties of the final casting products will be significantly affected negatively. When aluminium reacts with water in the atmosphere (like moisture), hydrogen gas is reduced and be dissolved in liquid aluminium. During the solidification, hydrogen solubility of the aluminium decreases and, released hydrogen results in the formation of pores. The porosity appears especially close to the surface in the form of small numbers of large gaps or small pores in the piece with a thick section [4, 5].

In this study, the porosity structures of AlSi9Cu3, AlSi12Cu1Fe and AlSi12Fe aluminium alloys were investigated before and after degassing. The porosity diameters and depths of the pieces were measured by computed tomography (CT). In addition, pore diameters and quantities were examined by means of a stereo microscopy (SM) method and the results were compared with the results which were obtained from the CT method.

II. EXPERIMENTAL STUDIES

Samples were taken before and after degassing for three different molten alloys. The alloys were AlSi9Cu3, AlSi12Cu1Fe and AlSi12Fe. In these experimental studies, a vacuum density tester (Mk 2200), which runs by the comparison of samples solidified under vacuum (80 mbar) and under atmospheric conditions (1000 mbar), was used to measure their hydrogen content. Thus, four different samples were casted from each alloy; before and after degassing and under vacuum and atmospheric conditions. Metkon Metacut 251 precise cutting device was used for sample cutting, YXLON MU2000-D device for CT imaging and Bestscope BUC2B-500C stereo microscope for micrographs were utilized.

III. RESULTS AND DISCUSSION

Dissolved hydrogen amounts in alloys were measured and shared in Table 1. The results were same for both atmospheric and vacuum conditions and, it proves the reliability of the analyses. Briefly it can be understood that conducting tests either under vacuum or under atmospheric conditions is viable to measure the dissolved hydrogen content accurately. Moreover, changing hydrogen amounts show that degassing was a successful operation to reduce the hydrogen content of the alloys to prevent the formation of porosities in the solidification of casting products.

Table 1: Hydrogen content of the alloys (wt.%).

Alloys	Before degassing	After degassing
AlSi9Cu3	6.37	1.50
AlSi12Cu1Fe	10.19	0.00
AlSi12Fe	13.85	1.53

CT and SM micrographs of AlSi9Cu3 alloy before degassing were given in Figure 1 for the solidified sample under vacuum, in Figure 2 for atmospheric conditions. Furthermore, Figure 3 and 4 show the samples which were solidified under vacuum and atmospheric conditions respectively after degassing. Reduced atmosphere made the pores visible and, positive effect of the degassing operation was also proved by the microstructure.

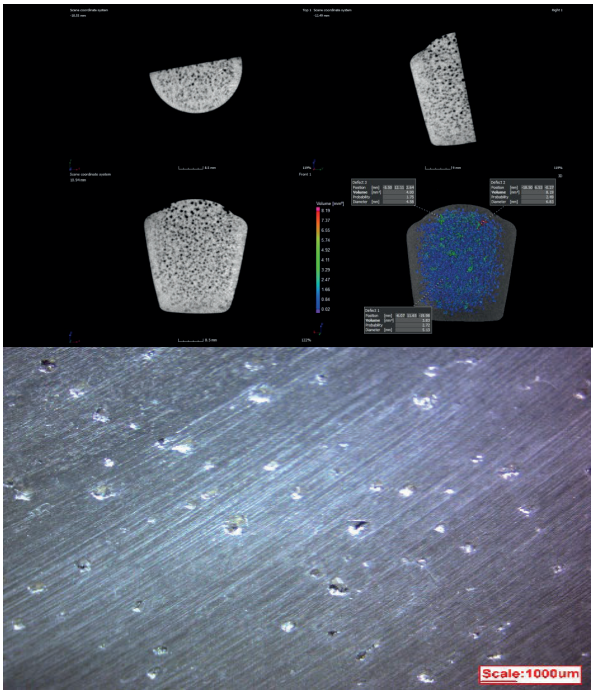


Figure 1: CT and SM images of AlSi9Cu3 before degassing for the solidified sample under vacuum.

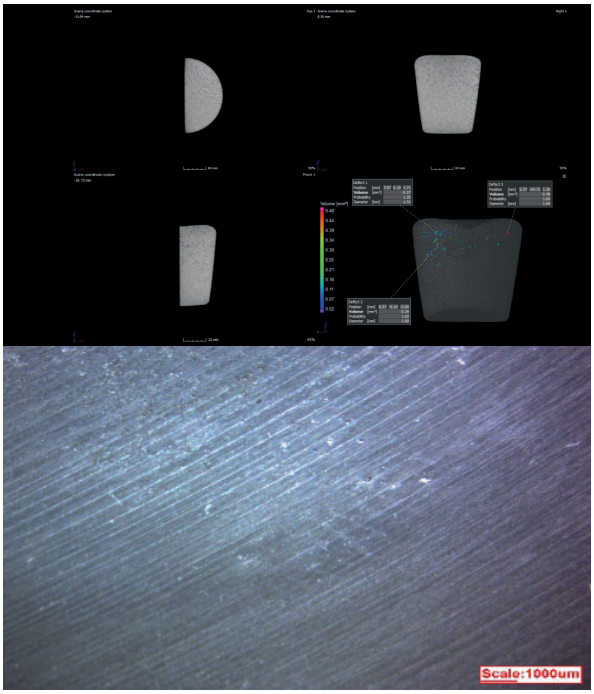


Figure 2: CT and SM images of AlSi9Cu3 before degassing for the solidified sample under atmospheric conditions.

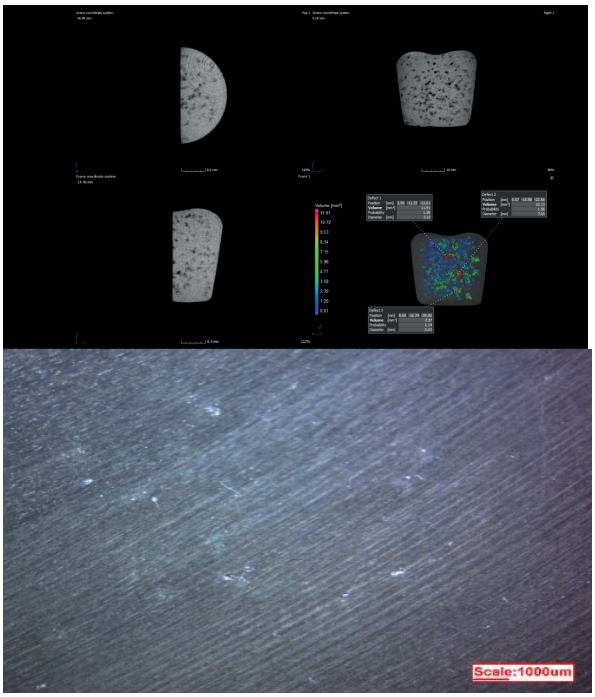


Figure 3: Figure 1: CT and SM images of AlSi9Cu3 after degassing for the solidified sample under vacuum.

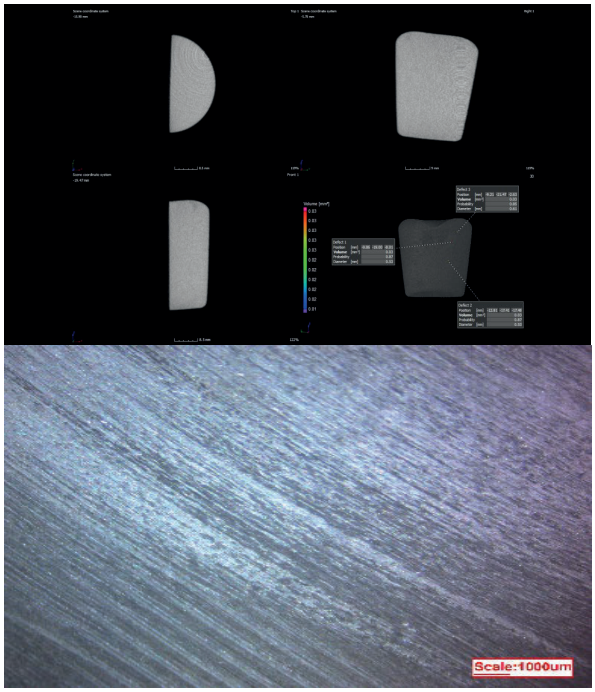


Figure 4: CT and SM images of AlSi9Cu3 after degassing for the solidified sample under atmospheric conditions.

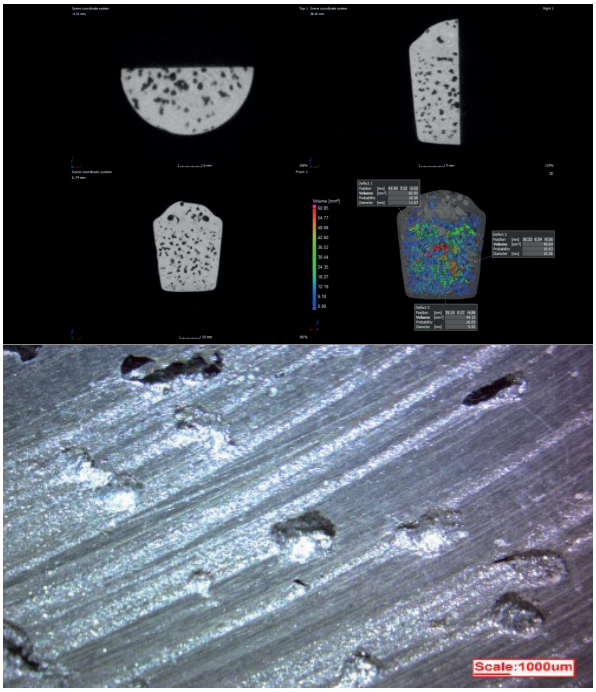


Figure 5: CT and SM images of AlSi12Cu1Fe before degassing for the solidified sample under vacuum.

From Figure 5 to Figure 8, the CT and SM micrographs of the tested samples were given before and after degassing operations for vacuum and for atmospheric conditions respectively. Pores were visible too in vacuum solidified samples in comparison to the samples which were solidified under atmospheric conditions. But, it does not mean that vacuum solidification results in the formation of pores in a huge ratio than that of the atmospheric conditions. With respect to the working mechanism of the hydrogen amount measurement device, it can be understood that under vacuum conditions dissolved hydrogen gas was released and caused the formation of the pores whereas the dissolved hydrogen was still in casting products for the samples solidified under atmospheric conditions. Those results were consistent with the results from Table 1 showing dissolved hydrogen amounts in the alloys.

In the last experimental series, the micrographs of the sample were taken for AlSi12Fe alloy for different solidification techniques before and after degassing operation. The results were similar to the previous test results from AlSi9Cu3 and AlSi12Cu1Fe alloys. Degassing was a successful method to remove the dissolved hydrogen content of the alloy (Figure 9-12).

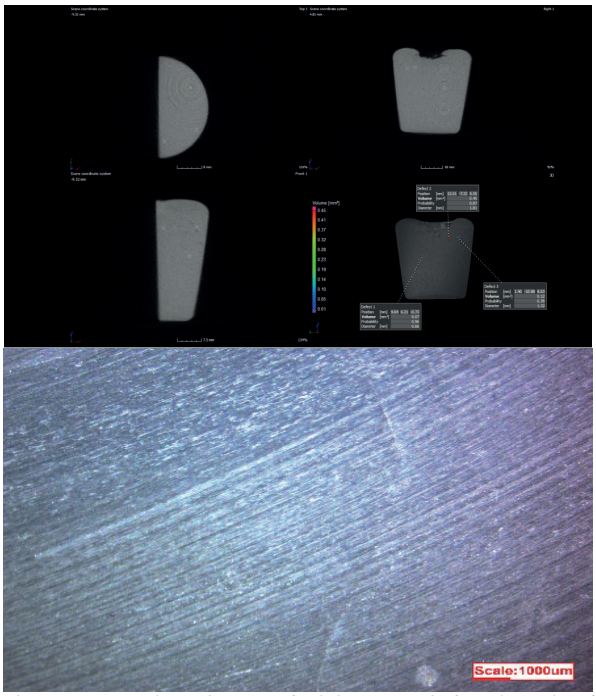


Figure 6: CT and SM images of AlSi12Cu1Fe before degassing for the solidified sample under atmospheric conditions.

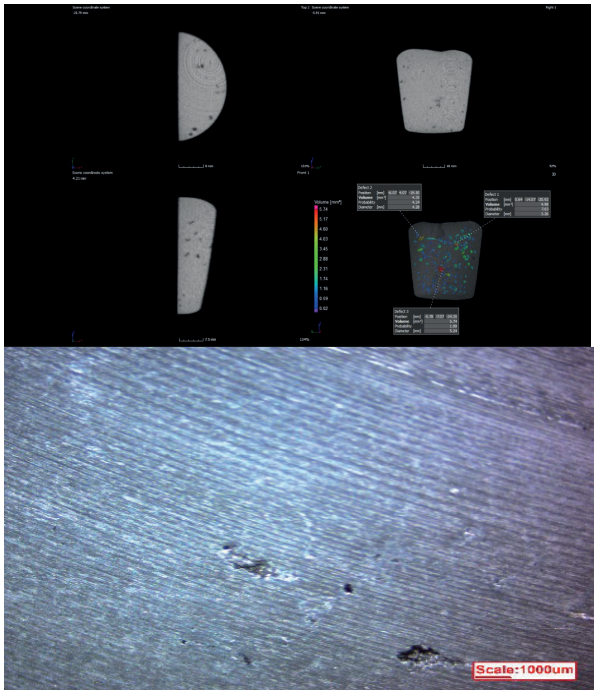


Figure 7: CT and SM images of AlSi12Cu1Fe after degassing for the solidified sample under vacuum.

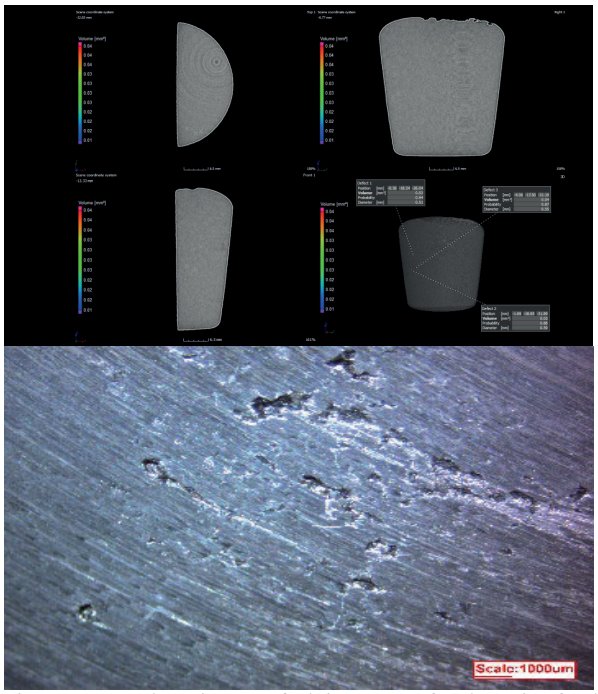


Figure 8: CT and SM images of AlSi12Cu1Fe after degassing for the solidified sample under atmospheric conditions.

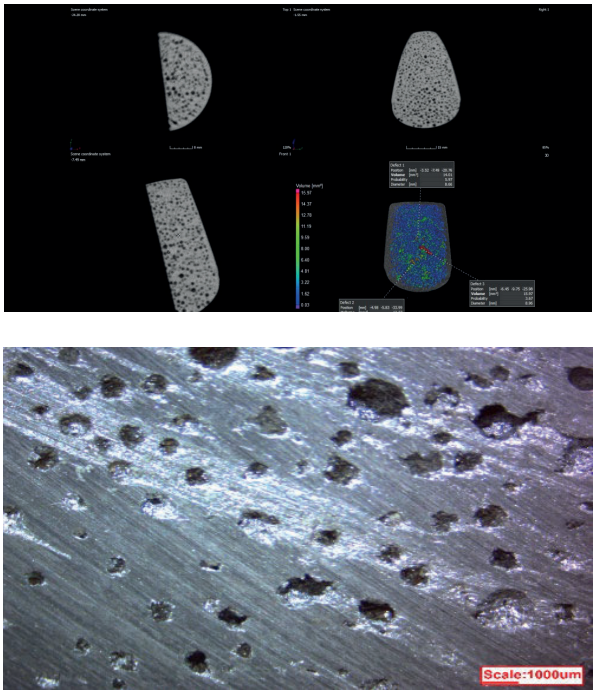


Figure 9: CT and SM images of AlSi12Fe before degassing for the solidified sample under vacuum.

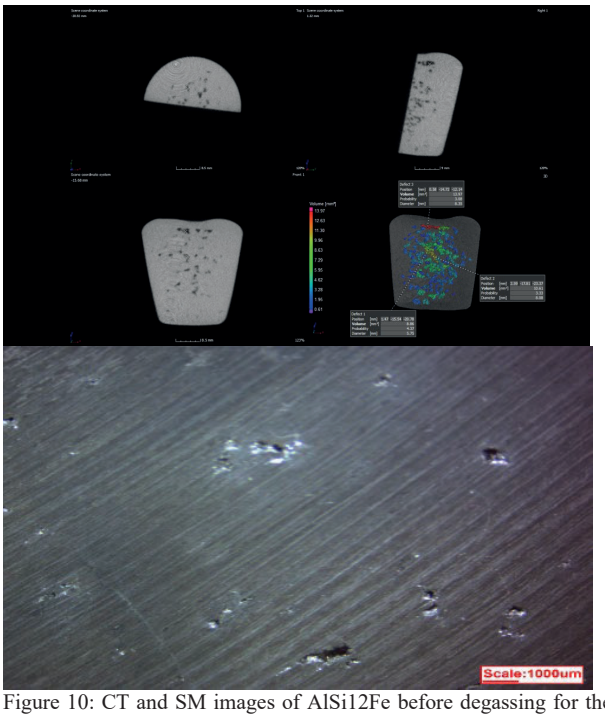


Figure 10: CT and SM images of AlSi12Fe before degassing for the solidified sample under atmospheric conditions.

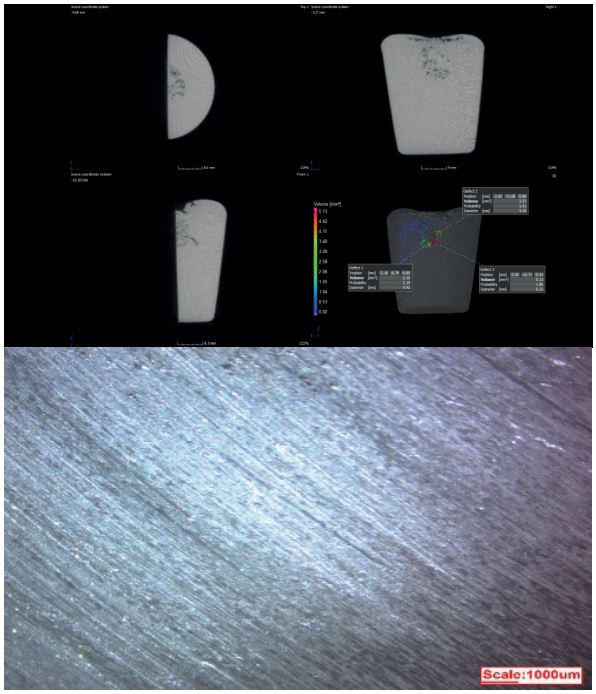


Figure 11: CT and SM images of AlSi12Fe after degassing for the solidified sample under vacuum.

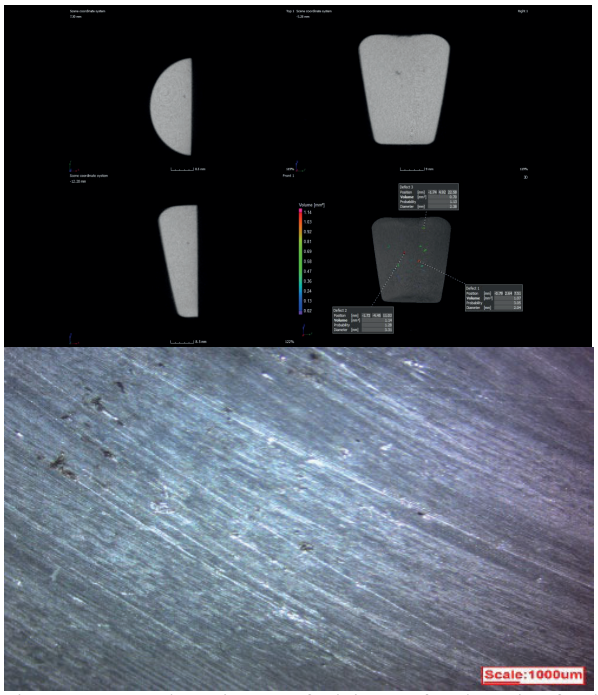


Figure 12: CT and SM images of AlSi12Fe after degassing for the solidified sample under atmospheric conditions.

IV. Conclusion

AlSi9Cu3, AlSi12Cu1Fe and AlSi12Fe aluminium high pressure die-casting (HPDC) alloys were investigated to understand their gas porosity formation behaviours before and after degassing operation under different solidification environments in a dissolved hydrogen measurement device by the comparison of their density change. The results were supported by means of CT and SM techniques by taking the micrographs of the samples. Degassing operations reduced the dissolved hydrogen amounts to the values less and equal to ~1.5%. Those values were acceptable for industrial practice to obtain sound HPDC products. This study is still ongoing and, authors plan to extend the experiments by using further characterizations of HPDC products which will be casted from the alloys which had used in the present study.

Acknowledgment

The authors are pleased to acknowledge the financial support for this research from Yalova University, Scientific Research Projects Department under grant number 2019/YL/0003

References

[1] International Aluminium Institute, http://www world-aluminium org/
[2] J R , Davis, ASM Specialty Handbook - Aluminum and Aluminum Alloys, ASM Int, 1993, 3–12
[3] Cinarli, U , Turan, A Investigation of Alumina-Based Ceramic Production from Aluminum Black Dross Mining, Metallurgy & Exploration (2020) https://doi org/10 1007/s42461-020-00344-0
[4] Ertürk S Ö Al Alaşımlarının Basınçlı Dökümünde Yolluk Sisteminden Kaynaklanan Gaz Problemlerinin Simülasyon Tekniği ile İncelenmesi, Sakarya Üniversitesi, Fen Bilimleri Enstitüsü, Metalurji ve Malzeme Mühendisliği Bölümü, Yüksek Lisans Tezi, 2010
[5] Patel J M , Pandya Y R ve Patel R C Optimization for Shrinkage Porosity on Aluminium Alloy ADC-12 Material, International Journal for Scientific research & Development Vol 5, Issue 09, ISSN: 2321-0613, 2017

Effect of Dopants on Akermanite/Hydroxyapatite Bioceramic Powders

B. BULUT[1], Ş. DUMAN[2], M. EROĞLU[3], M.G. ÖZGÜVEN[4]

[1] Bursa Technical University, Bursa/Turkey, mbusrabulut@gmail.com
[2] Bursa Technical University, Bursa/Turkey, seyma.duman@btu.edu.tr
[3] Bursa Technical University, Bursa/Turkey, murat.eroglu@btu.edu.tr
[4] Istanbul Technical University, Istanbul/Turkey, gultekinmi@itu.edu.tr

***Abstract* - This work reports about iron oxide and zinc oxide doped akermanite/hydroxyapatite composite powders produced by mixing in an aqueous solution of oxide precursor powders. Iron oxide and zinc oxide were used as dopants in different concentrations between 1-7%. The effect of dopants on the powders was investigated utilizing the structural and morphological characterization techniques such as X-ray diffraction (XRD) analysis, scanning electron microscopy coupled with energy dispersive X-ray spectrometry (SEM-EDS), particle size analysis (PSA). The XRD patterns showed that akermanite was successfully obtained in the predominant crystal phase. The iron oxide and zinc oxide additions showed a slight influence on the powder morphology. Furthermore, the morphologies of the composite powders were affected by the amount of dopant.**

Keywords - **Akermanite, Hydroxyapatite, Iron oxide, Zinc oxide.**

I. Introduction

Bioceramics, especially calcium phosphate-based ceramics, attracted much attention for biomedical applications due to their biocompatibility, bioactivity, and osteoconductive properties [1-3]. *Hydroxyapatite (HA, $Ca_{10}(PO_4)_6(OH)_2$)* is one of the most widely used calcium phosphate-based bioceramics with its outstanding biological responses to the physiological environment [4]. However, disadvantages such as being not biodegradable and low fracture toughness limit their wider application [5].

Recently, *Akermanite (AK, $Ca_2MgSi_2O_7$)* as a novel bioceramic has been the focus of researchers with its superior degradation rate, its biodegradability, and its controllability of mechanical properties [1, 6-8]. Furthermore, it is known that AK has a stimulating effect on the body with the inorganic ions it contains, and this stimulating effect is related to its bioactivity [9, 10]. The combined use of these materials can overcome the limitations of their independent use. It is also known that the physical and biological properties of these materials can be significantly improved by including metal oxides [11]. Among them, *iron oxide (Fe_2O_3)* and *zinc oxide (ZnO)* stand out by showing biocompatibility and antibacterial activity [12, 13].

In this study, the effect of different amounts of iron oxide and zinc oxide on the structural and morphological properties of akermanite/hydroxyapatite bioceramic powders was investigated. It was shown that increasing the amount of dopant change the structural morphology of the bioceramic composites.

II Experimental

A. Materials

Synthetic hydroxyapatite powders (reagent grade, molecular weight 502.31, 10-20 µm particle size range) were purchased from Sigma Aldrich™, USA. Elemental Fe_2O_3 and ZnO powders (ZnO, 99.9% purity) was obtained from Alfa Aesar™, USA.

B. Powder Preparation

AK and HA powders, with an equal proportion for 0,4 g of total powders were mixed by dissolving in a 10 ml aqueous solution. Hereafter, Fe_2O_3 and ZnO powders at different weight percentages (1, 3, 5, and 7 wt%) were added into the obtained homogeneous solutions and stirred at room temperature for 30 minutes. Prepared eight different solutions were dried in a laboratory oven and doped powders were obtained. Flow chart summarized the experimental details is given in Figure 1 and applied sample compositions and their experimental codes used in this study are given in Table 1.

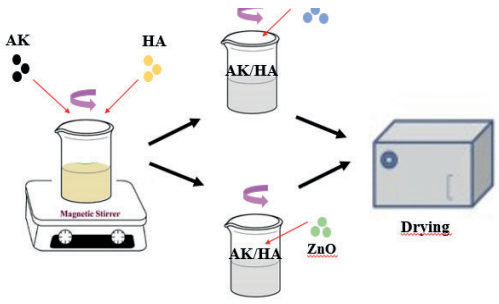


Figure 1: Schematics of the preparation process.

Table 1: Sample compositions and their experimental codes

Sample compositions	*Codes*
AK/HA %1 wt. Fe_2O_3	1F-AK-HA
AK/HA %3 wt. Fe_2O_3	3F-AK-HA
AK/HA %5 wt. Fe_2O_3	5F-AK-HA
AK/HA %7 wt. Fe_2O_3	7F-AK-HA
AK/HA %1 wt. ZnO	1Z-AK-HA
AK/HA %3 wt. ZnO	3Z-AK-HA
AK/HA %5 wt. ZnO	5Z-AK-HA
AK/HA %7 wt. ZnO	7Z-AK-HA

C. Characterization

X-ray diffraction (XRD) pattern data were determined at room temperature using powder samples with CuK_α radiation (λ = 0.1540 nm) by Bruker™ Discovery D8 Advantage XRD. Microstructural analysis of the Fe_2O_3 and ZnO doped AK/HA composite powders were observed by using Carl Zeiss™ Gemini 300 scanning electron microscope (SEM, operated at 20 kV) coupled with an energy-dispersive X-ray spectrometer (EDS) after they were sputter-coated with gold-palladium. Particle size measurements of the powders were conducted using Malvern™ Mastersizer 3000 particle size analyzer.

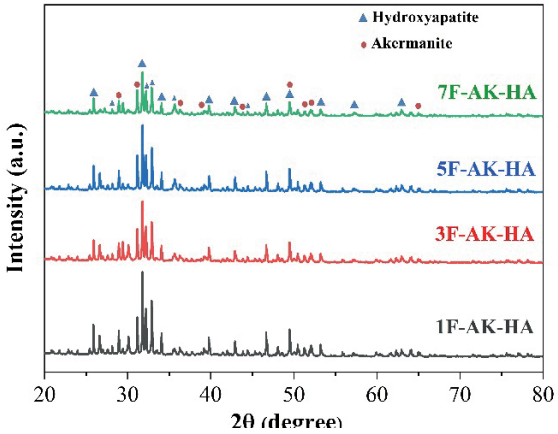
III. Results and Discussion

Akermanite powders were synthesized in our previous study [14] by spray drying and subsequent heat treatment at 1300 ºC dual methods used. In this study, the composite powders were successfully prepared by mixing in an aqueous solution of Fe_2O_3, ZnO, the synthesized AK powders, and synthetic HA powders.

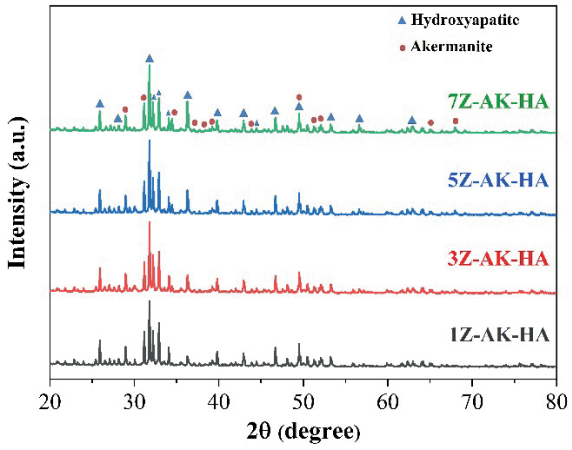
Figure 2(a)–(b) presents the XRD diffraction patterns of the AK/HA composite powders prepared by adding Fe_2O_3 and ZnO between 1 and 7 wt %. As can be seen in Fig. 2(a), AK [14] and HA [15] phases were detected in the XRD of the Fe_2O_3 doped AK/HA composite powders. As can be seen in Figure 2(a), hydroxyapatite (JCPDS# 64-0738) with a hexagonal structure is the dominant phase and a small amount of akermanite (JCPDS# 35-0592) with a tetragonal structure could also be detected in the XRD of the Fe_2O_3 doped AK/HA composite powders. The increasing amount of the Fe_2O_3 dopant resulted in a decrease in the intensity of related hydroxyapatite and akermanite peaks. Moreover, no impurity phases and Fe_2O_3 are detected in the resultant composite powders.

Figure 2(b) shows an XRD pattern of the ZnO doped AK/HA composite powders prepared by mixing in an aqueous solution. Again, hydroxyapatite and akermanite phases exist in the composite powders. ZnO peaks are not seen in the resultant composite powders. Furthermore, no significant peak broadening and peak intensity were observed with an increasing amount of the ZnO dopant (Figure 2(b)).

The SEM micrographs of all the Fe_2O_3 doped AK/HA composite powders and the ZnO doped AK/HA composite powders are shown in Figures 3 and 4, respectively. As can be seen from Figures, preparation by mixing in an aqueous solution of the composite powders led to comprise of lamellar shaped particles. It was observed that the morphology of composite powders slightly changed with the addition of Fe_2O_3 or ZnO from 1 wt. % to 7 wt. % and had similar particle sizes. Figures 3 and 4 present similar microstructures, figures exhibit a glassier structure, which is probably caused by the akermanite content formerly detected in the XRD pattern in Fig. 2.



(a)



(b)

Figure 2: XRD patterns of the composite powders: a) Fe_2O_3 doped AK/HA and b) ZnO doped AK/HA.

Table 1: Sample compositions and their experimental codes used in this study.

Sample Codes	*Particle Sizes (PS; µm)*	*Specific Surface Areas (SSA, m^2/g)*
1F-AK-HA	10,1	7,97
3F-AK-HA	12,2	6,41
5F-AK-HA	11,8	6,89
7F-AK-HA	11,5	7,06
1Z-AK-HA	11,5	6,97
3Z-AK-HA	12,9	6,33
5Z-AK-HA	10,8	7,47
7Z-AK-HA	10,7	7,51

The particle sizes and specific surface areas of Fe_2O_3 and ZnO doped AK/HA composite powders are given in Table 2.

The particle sizes for Fe_2O_3 doped AK/HA composite powders were between the range of 10.1 – 12.2 μm. It was observed that with Fe_2O_3 doping and the increase of dopant concentrations, particle size was not significantly changed. The specific surface

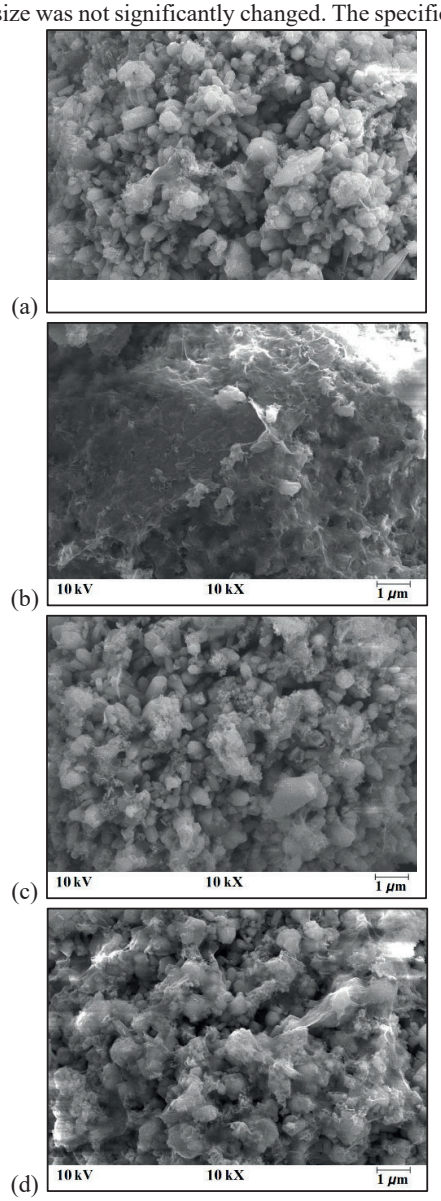


Figure 3: SEM micrographs of the Fe_2O_3 doped AK/HA powders: a) 1F-AK-HA, b) 3F-AK-HA c) 5F-AK-HA d) 7F-AK-HA.

area values were in the range of 6.41 - 7.97 m^2/g. While the particle sizes for ZnO doped AK/HA composite powders were between the range of 10.7 – 12.9 μm, the specific surface area values were in the range of 6.33 - 7.51 m^2/g. Similarly, it was interpreted that the particle sizes were not dramatically changed with ZnO doping and the increase of dopant concentrations. It has been determined that the type of dopant and the concentration of the dopant do not affect the physical properties such as particle size and specific surface area.

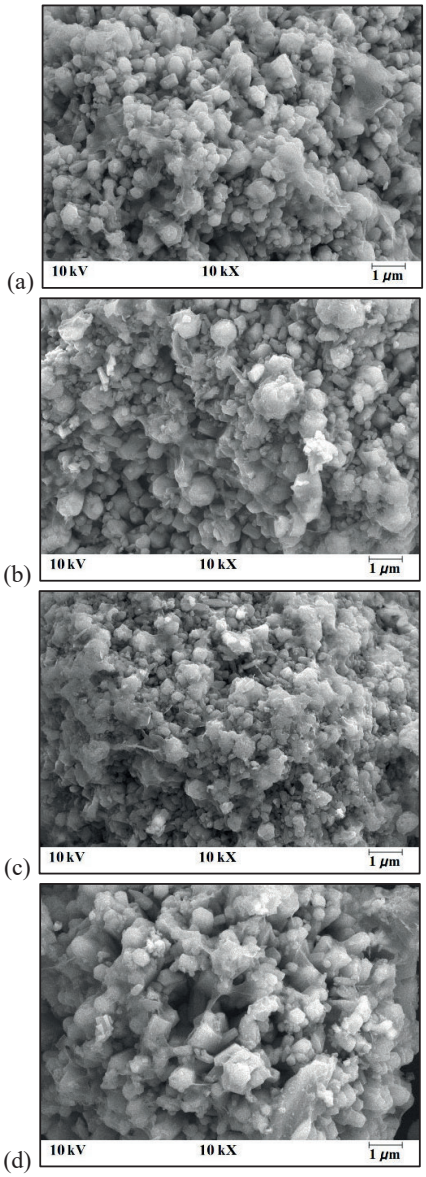


Figure 4: SEM micrographs of the ZnO doped AK/HA powders: a) 1Z-AK-HA b) 3Z-AK-HA c) 5Z-AK-HA d) 7Z-AK-HA.

Figure 5 revealed that particle size distributions of Fe_2O_3 and ZnO doped AK/HA composite powders have similar peak intensities. While the peak of AK/HA composite powders slightly shifted towards the higher particle size value by increasing the amount of Fe_2O_3 dopant, it shifted towards lower for the increasing amount of ZnO dopant.

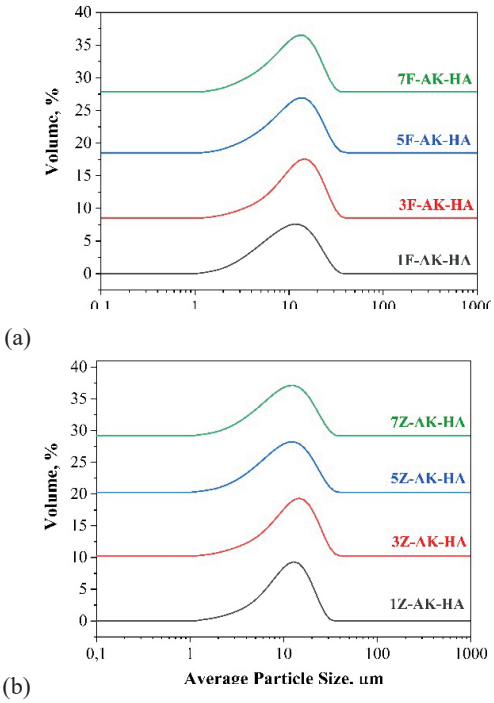


Figure 5: The particle size distributions of Fe_2O_3 and ZnO doped AK/HA powders: a) 1Z-AK-HA b) 3Z-AK-HA c) 5Z-AK-HA d) 7Z-AK-HA.

IV. Conclusions

Fe_2O_3 and ZnO doped AK/HA composite powders were successfully prepared by mixing in an aqueous solution of oxide precursor powders. AK/HA composite powders prepared by adding Fe_2O_3 and ZnO between 1 and 7 wt. % were indicated by XRD, SEM, and particle size distribution characterizations. According to the XRD analysis, hydroxyapatite and akermanite were successfully obtained in the predominant crystal phase in the composite powders. Fe_2O_3 and ZnO dopants slightly showed an influence on the structure morphology.

Appendix

There is no appendix in this study.

Acknowledgment

The authors thank Bursa Technical University Central Research Laboratory for providing laboratory facilities for this research.

Reference

[1] Huang, Y , et al , *In vitro and in vivo evaluation of akermanite bioceramics for bone regeneration.* Biomaterials, 2009 **30**(28): p 5041-5048

[2] Ragel, C , M Vallet-Regı, and L Rodrıguez-Lorenzo, *Preparation and in vitro bioactivity of hydroxyapatite/solgel glass biphasic material.* Biomaterials, 2002 **23**(8): p 1865-1872

[3] LeGeros, R Z , *Properties of osteoconductive biomaterials: calcium phosphates.* A Publication of The Association of Bone and Joint Surgeons®| CORR®, 2002 **395**: p 81-98

[4] Kalita, S J and H A Bhatt, *Nanocrystalline hydroxyapatite doped with magnesium and zinc: Synthesis and characterization.* Materials Science and Engineering: C, 2007 **27**(4): p 837-848

[5] Wu, C and J Chang, *A review of bioactive silicate ceramics.* Biomedical materials, 2013 **8**(3): p 032001

[6] Wu, C and J Chang, *Degradation, bioactivity, and cytocompatibility of diopside, akermanite, and bredigite ceramics.* Journal of Biomedical Materials Research Part B: Applied Biomaterials: An Official Journal of The Society for Biomaterials, The Japanese Society for Biomaterials, and The Australian Society for Biomaterials and the Korean Society for Biomaterials, 2007 **83**(1): p 153-160

[7] Hou, X , et al , *Effect of akermanite morphology on precipitation of bone-like apatite.* Applied Surface Science, 2011 **257**(8): p 3417-3422

[8] Kokubo, T , *Bioactive glass ceramics: properties and applications.* Biomaterials, 1991 **12**(2): p 155-163

[9] Xynos, I D , et al , *Ionic products of bioactive glass dissolution increase proliferation of human osteoblasts and induce insulin-like growth factor II mRNA expression and protein synthesis.* Biochemical and biophysical research communications, 2000 **276**(2): p 461-465

[10] Wu, C , et al , *In vitro bioactivity of akermanite ceramics.* Journal of Biomedical Materials Research Part A: An Official Journal of The Society for Biomaterials, The Japanese Society for Biomaterials, and The Australian Society for Biomaterials and the Korean Society for Biomaterials, 2006 **76**(1): p 73-80

[11] Feng, P , et al , *Toughening and strengthening mechanisms of porous akermanite scaffolds reinforced with nano-titania.* RSC advances, 2015 **5**(5): p 3498-3507

[12] Gnaneshwar, P V , et al , *Ramification of zinc oxide doped hydroxyapatite biocomposites for the mineralization of osteoblasts.* Materials Science and Engineering: C, 2019 **96**: p 337-346

[13] Sneha, M and N M Sundaram, *Preparation and characterization of an iron oxide-hydroxyapatite nanocomposite for potential bone cancer therapy.* International journal of nanomedicine, 2015 **10**(Suppl 1): p 99

[14] Bulut, B , *Kemik doku uygulamaları için TiO2/akermanit katkılı kitosan doku iskelelerinin üretimi ve karakterizasyonu.* Bursa Teknik Üniversitesi/Fen Bilimleri Enstitüsü, 2019

[15] Sombra, A and C Silva, *Structural and Mechanical Study of the Sintering Effect in Hydroxyapatite Doped with Iron Oxide.* 2008

A Gas Sensor Design and Simulation with ZnO and TiO_2

Gozde Konuk Ege [1], Huseyın Yuce[2], and Garıp Genc[2]

[1] [2] Mechatronics Program, Istanbul Gedik University, Istanbul, 34913, Turkey

Abstract **- Micro Electro-Mechanical System (MEMS) based devices offer innovative approaches in sensor technologies with the advantages of high efficiency and miniaturization. The most important stage in the development of new generation MEMS-based devices is the design and optimization stage. However, device design and optimization processes are developed in a laboratory by empirical approaches. This causes time loss and creates an unnecessary waste of resources. In this study, it is aimed to design and analyze two gas sensors based on ZnO and TiO_2 sensing layers. Electro-thermal analysis of the sensor structure was carried out at room temperature and high temperature (294,15K-573,15K) and heat transfer parameters were compared. According to the simulation results, it is obtained that, as the applied temperature increases to the sensor, the temperature over the sensing layer increases linearly. It is compatible with the literature. The temperature on the ZnO surface increases to three times the TiO_2 surface temperature. The heat transfer results obtained will be used as a guide for device design and optimization in future works. In this way, as a result of numerical analysis, a MEMS-based device will be produced with high accuracy. Thus, time and resources will be saved.**

Keywords **- Gas sensor; sensing layer; finite element method (FEM); heat transfer**

I. INTRODUCTION

apidly increasing environmental pollution and consequently increasing health problems have become one of the most important problems of today. In addition, different gases are frequently used in the manufacturing industry. This situation directly affects to human health. The integration of smart devices developed for the benefit of humanity into our daily life is becoming more and more popular with decades due to their usefulness. Gains in MEMS technology also has been affected sensor technologies. Sensors produced in line with micro technologies have advantages in terms of cost, miniaturization, and sensitivity as smart sensors. In this view, the development of gas detection devices is important for widespread uses, such as monitoring gas accumulation, medical diagnostics, food quality assurance and the safety of industrial processes or domestic systems. Today, in parallel with the technological progress in nano-electronic structuring and the point it has reached, gas sensor technologies continue to progress with similar acceleration [1–5].

Micro/nano technology is a multi-disciplinary technology because of using material science, structure-control design and finite-element method (FEM). Material technology allows the use of new generation materials and semiconductor technologies in miniature device development. Moreover, structural/control design and analysis methods offer high accuracy and analysis advantages with high accuracy.

It was observed for the first time in 1953 that the conductivities of semiconductor materials are changed by gas absorption [6]. Subsequently, the first semiconductor gas sensor patented in 1962 [7]. Afterwards, gas-sensing devices are researched and examined worldwide. Researches on gas detection mechanisms is focused on designing gas detection elements that can detect gases harmful to human health and nature with high performance detection materials as a consequence various gas detection technologies have been developed for use in gas detection processes [8–13].

Table 1: Gas detection methods used in gas detection processes

Gas sensors related to detecting material	Gas sensors related to detecting mechanism
Metal oxide semiconductors	Resistance Change
Carbon nanotubes	Optical methods
Polymers	Acoustic methods
	Gas chromatography
	Calorimetric methods

Gas detection processes can be examined in two classes, according to the sensing material and the sensing mechanism, as shown in Table 1. In gas sensors; Metal oxide semiconductors, carbon nanotubes and polymers are used as the material in the gas detection layer. On the other hand, resistance changing, optical methods, acoustic methods (eg; SAW sensor), calorimetric method and gas chromatography can be used as the detection mechanism. Metal oxide-based gas sensors attract researchers' attention thanks to their high detection capability, repeatability and simple manufacturing techniques. MO_X-based

gas sensors are used to sense NH_3, H_2, HCl, NO_2, H_2S, CO, CO_2, SO_2, O_3, Ar, N_2, VOCs, LPG gases [14].

NH_3, H_2, HCl, NO_2, H_2S, CO, CO_2 gases threat human health depending on their conditions of use and release rates of these gases. Table 2 presents that the areas of use or release of these gases and their danger. For example; NH3 gas appears in the chemical, food and health industry and causes a toxic effect. NO, HCl, CO_2 gases cause respiratory disorders. The release of high levels of H_2S and CO, H_2 gases can result in death.

Table 2. According to the encountered area, and their dangers various hazardous gases

Hazardous Gases	The encountered Area	Dangers
NH_3 [15]	Chemical and food industry, medical diagnosis	Toxic effect, vomiting, headache
NO [16–19]	Combustion of fossil fuels, power plants and vehicle engines	Respiratory diseases, asthma
NO_2 [20]	Automobile exhaust fumes, nitric acid production, coal and fuel combustion	Acid rain, environmental pollution
H_2 [21–25]	Renewable energy sources, transport systems and biomedical devices	Explosive, flammable
H_2S [26–29]	Oil/natural gas, geothermal energy and bacterial decomposition	Occupational diseases, death in high concentration
HCl [30]	Semiconductor and chemical industry	Toxic effect, skin burns, respiratory disorders
CO [31]	Fire events	Toxic effect , death
CO_2 [32]	Cellular respiration and burning of fossil fuels	Respiratory disorders, explosion, Global warming, explosion

The working principle of MO_X gas sensors are usually based on the resistivity change of the semiconductor, depending on the gas absorption. When gas analytes come into contact with the sensing layer, the adhesion of oxygen atoms occur on the surface and semiconductor sensing material reacted to the gas analytes. In this way, sensing layer resistivity is decreased according to the analyte concentration. However, the worst feature of MOx-based gas sensors is that they have a high operating temperature. This situation requires high power consumption and an external heat source. Therefore, many studies are carried out to obtain high operating performance at low temperatures. In this study, two MOx-based gas sensors model are designed. One of the gas sensors consists of TiO_2 as a sensing layer and another is consist of ZnO. Three-dimensional design and simulation are done by using SolidWorks and COMSOL Multiphysics software. The electro-thermal analysis is done to observe the surface temperature and electrical dispersion along the sensing layer. The effect of TiO_2 and ZnO semiconductors heat transfer is evaluated over the sensing layer at room temperature and different high temperatures.

II. STRUCTURE DESIGN

Gas sensors are smart devices used to sense many other gases with concentrations ranging from ppm and ppb in the presence of other low concentrations gases [33]. Although it is considered to be an apparently simple operating principle, the gas detection mechanism is quite complicated. MEMS-based gas sensors consist of two main layers as shown in Fig.1. Gas sensing performance and especially sensitivity are controlled these two layers. A chemical or biochemical interface is required to obtain sufficient sensitivity and selectivity for the analyte. This interface is called the sensing layer, this is the top layer. Gas detection mechanism is related to the change of one or more physical properties of the sensing material such as mass, elastic hardness, viscosity, electrical conductivity, and electrical permeability. The conductivity of the sensing material is expected to change according to the gas concentration. It is also desired that this change in conductivity be reversible. The performance of a gas sensor depends on the sensor configuration as well as the interaction of the sensing material with the gas. It is important to understand the interaction between the sensor material and the analyte molecules of basic sensor parameters such as sensitivity and selectivity. If this interaction is weak; sensitivity and selectivity are weakened when the sensor shows good recycling or recycling is weakened while showing good sensitivity and selectivity [34]. Another layer is the transducer layer. This layer converts the sensing analyte signals from the chemical interaction occurring in the sensing layer into an electrical signal.

Fig. 1 MEMS-based gas sensors layer structure

Sensor specifications, for instance sensitivity, selectivity, response and recovery times, are affected by the properties of the sensing material. In this study, TiO_2 and ZnO are preferred as a sensing material. ZnO has high chemical and thermal stability, piezoelectric properties. These provide excellent detection response. It has a low cost. It offers easy production and non-toxic. Thanks to these advantages, they are preferred in gas sensor applications [1]. TiO_2 is also stable, low cost and

non-toxic semiconductor material. It has unique electrical, optical and catalytic gas sensing material because of its wide bandgap. In smart device applications, the band structure provides superior electronic and optical properties [35]. The designed gas sensors structure is presented in Fig.2.

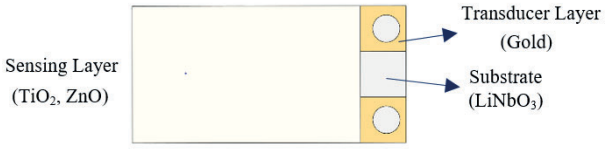


Fig. 2. Gas sensor structure

$LiNbO_3$ wafer is used as a substrate. On the substrate, gold electrodes are evaporated in a vacuum for electrical conductivity. Above the gold electrodes, the sensing layer is sputtered. Mechanical properties are given in Table 3. The hot plate to be used as a high heat source will be placed where the gas sensor socket in the gas detection chamber.

Table 3. Mechanical properties of sensor structure

Material	Size		Thermal Conductivity (W/(m K))	Density (kg/m3)	Young's Modulus (GPa)
	Length x Wide (mm)	Thickness (mm)			
LiNbO3 [36]	6x3	0 5	5 6	4640	170
Gold [37]	6x1	1x10-4	317	19300	70
TiO_2 [38]	5x3	5x10-5	8 9	4230	230
ZnO [39]	5x3	5 x10-5	60	5676	210

III. ELECTRICAL CONDUCTIVITY AND HEAT TRANSFER ANALYSIS

The heat transfer analysis and joules heating simulation are important for observing the homogeneous heat dispersion on the surface. Homogeneous heat dispersion is one of the determining factors of sensor sensitivity. Seeing that metal oxide gas sensors work according to the resistance change principle, it is expected to be homogeneous and constant at the electrical voltage on the sensing surface. For electrical conductivity simulation, the electrical field dispersion (E) is determined using Eq.1.;

$$E = -\nabla v \tag{1}$$

The electrical resistivity of most materials changes with temperature. To calculate electrical resistivity, a linear approximation is typically used. For electrical resistivity, (ρ) in certain temperature is determined by Eq.2;

$$\rho(T) = \rho_0(1 + \alpha(T - T_0)) \tag{2}$$

In Eq.2; ρ(T) is electrical resistivity in a certain temperature, ρ0 is electrical resistivity at 20 ºC, α is temperature coefficient at 20 ºC, T0 is 20 ºC and T is a certain temperature.

Acording to the electric field (E); electrical conductivity of the sensing material,(σ), current density (J) and joule heating power (Qj) are calculated by Eq.3., and Eq.4.;

$$J = \sigma E \tag{3}$$

$$Qj = \nabla J \tag{4}$$

Moreover, the heat transfer is calculated with Eq.5. The heat transfer depends on the heat flux, thermal conductivity, absolute depends on the thermal conductivity, temperature difference, and length of the body. In this equation, q, Δt, and k represent the heat flux, the temperature difference, and the thermal conductivity respectively [40] .

$$q = \frac{k.\Delta t}{L} \tag{5}$$

IV. RESULTS AND DISCUSSION

A. Electrical Dispersion

The sensing layer electrical dispersion of the gas sensor from room temperature to 573,15K is analyzed. A 9V initial voltage is applied to the gas sensor. In accordance with the resistivity of the sensor material, the average voltage above the sensing layer surface is observed and it is presented in Fig.3. At the ZnO surface, electrical dispersion has three different distinct regions.

In region I, the average electrical potential is observed steadily from room temperature to 423,15K. ZnO exhibits a negative temperature coefficient of resistance (NTCR) behavior in this region. At second region from 423,15K to 523,15K ZnO indicates a positive temperature coefficient of resistance (PTCR) behavior.

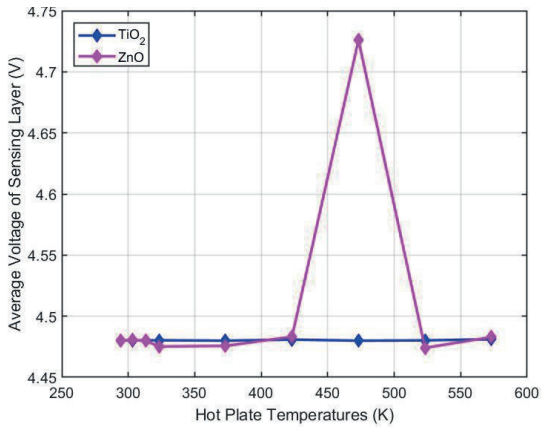


Fig.3. Average voltage above the sensing layer (TiO_2, ZnO)

Resistivity level is the lowest in this region. In region III from 523,15K to 573,15K resistivity increases again and ZnO shows a negative temperature coefficient in this region. There are several studies about the temperature-dependent resistivity of ZnO according to NTCR and PTCR behavior.

B. Temperature Dispersion

In this simulation; absolute pressure is 1 atm. First of all, the room temperature condition is done and heat dispersion is observed. Then, high temperature conditions (323,15K, 373,15K, 423,15K, 473.15K, 523,15K, 573,15K) are created by hot plate respectively. An initial temperature (T_0=room temperature) is 294,15K and the external temperatures (T_{ext}=hot plate temperature) were applied by increasing the temperature step by step from 323,15K to 573,15K. Fig.4, and Fig.5 shows the temperature profiles of TiO_2 and ZnO sensing layers with a base substrate.

A uniform sensing layer temperature is a requirement for sensitive operation of the sensor. Figure 4 and 5 shows uniform temperature dispersion on the sensing surface at different operating temperatures. Furthermore, Fig.6. shows the comparison of surface temperature at constant electrical voltage 9V. While the temperature on the TiO_2 surface at 473,15K was 489.56 K and 569.86 K on the ZnO surface, the temperature on the TiO_2 surface at 573,15K was 603.41 and 715.51 K on the ZnO surface.

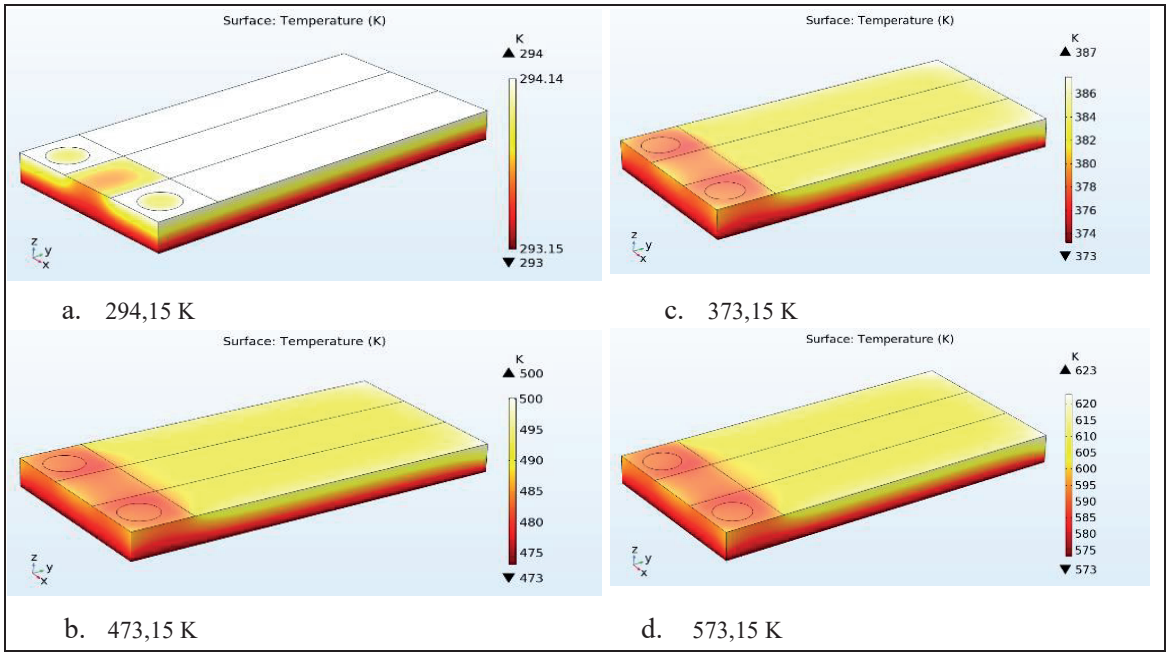


Fig.4. Temperature dispersion in TiO_2 layer, for the temperature 294,15K (a), 373,15K (b), 473,15K (c), 573,15 K (d)

The simulation results obtained are compatible with the literature [41], [42]. On the other hand, the TiO_2 surface electrical dispersion is observed steadily. TiO_2 shows a negative temperature coefficient from room temperature to 573,15K.

The comparison in Fig.6. depicts that sensing surface temperature increase on both surfaces; increased proportionally from the room temperature to 423.15 K.

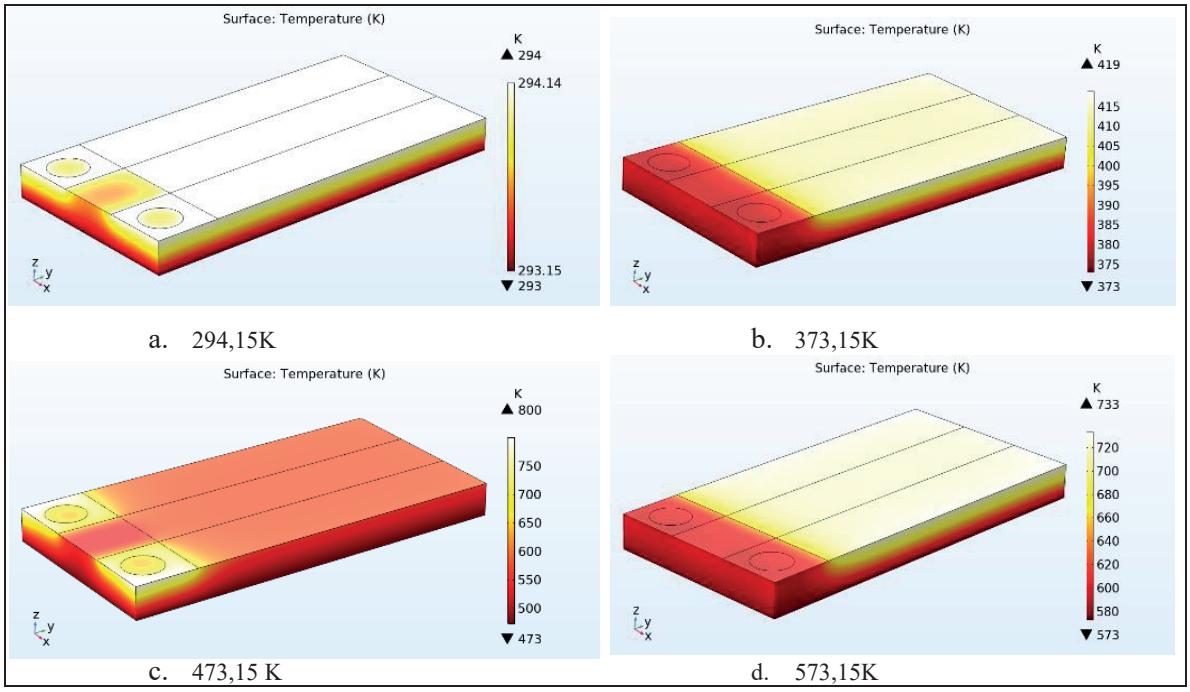


Fig.5. Temperature dispersion in ZnO layer, for the temperature 294,15K (a), 373,15K (b), 473,15K (c), 573,15K (d)

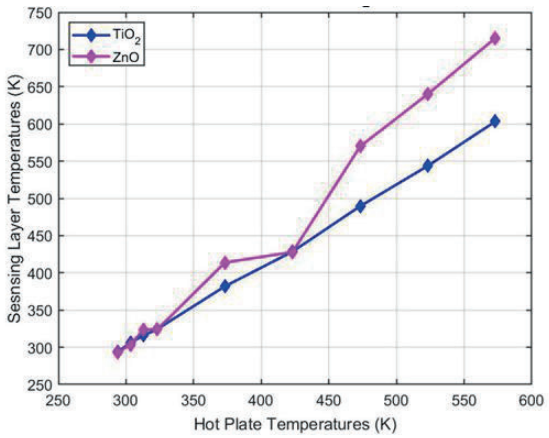


Fig.6. Sensing layer temperature at different hot plate temperatures

When 423.15 K exceeds; the temperature increase on ZnO surface gained acceleration compared to TiO_2 surface. While the temperature on the TiO_2 surface at 473,15K was 489.56 K and 569.86 K on the ZnO surface, the temperature on the TiO_2 surface at 573,15K was 603.41 and 715.51 K on the ZnO surface.

According to the simulation results, the optimum operating temperature was determined of ZnO and TiO_2 thin films. Sensing surface temperature rises above at the 473.15K, 523.15K, 573.15K operating temperature.

Consequently, heating inevitably occurs on the surfaces. In this case, the optimum working temperature for both thin-film structures with a thickness of 0.5 µm was determined as 423.15 K.

V. Conclusion

In this study, heat transfer analysis of the MOx-based gas sensor is discussed. Two sensing layer structures that used TiO_2 and ZnO as a material are designed and simulated. To achieve uniform heat dispersion above TiO_2 and ZnO sensing layer at room temperature and high temperature are analyzed by using FEM with COMSOL Multiphysics. The temperature profiles of sensing layers are analyzed and uniform heat dispersion was achieved in both thin films. As stated in simulation results, an optimum working temperature was determined for ZnO and TiO_2 thin films.

Our future works will be focused on the designed metal-oxide/polymer-based gas sensor at room temperature and high temperature and compared simulation results between them. Additionally, we aim to design and fabricate a gas sensor using metal oxide and metal oxide/polymer sensing layer.

ACKNOWLEDGMENT

This work was supported by Scientific Research Projects Coordination Unit of Istanbul Gedik University. Project number "GDK201905-14"

REFERENCES

[1] V S Bhati, M Hojamberdiev, and M Kumar, "Enhanced sensing performance of ZnO nanostructures-based gas sensors: A review," *Energy Reports*, no xxxx, 2019

[2] K H Rahman and A K Kar, "Titanium-di-oxide (TiO2) concentration-dependent optical and morphological properties of PAni-TiO2 nanocomposite," *Mater. Sci. Semicond. Process.*, vol 105, no April 2019, p 104745, 2020

[3] R Kumar and R Ghosh, "Selective determination of ammonia, ethanol and acetone by reduced graphene oxide based gas sensors at room temperature," *Sens. Bio-Sensing Res.*, vol 28, no January, p 100336, 2020

[4] L Kumar, I Rawal, A Kaur, and S Annapoorni, "Flexible room temperature ammonia sensor based on polyaniline," *Sensors Actuators, B Chem.*, vol 240, pp 408–416, 2017

[5] B Mondal, M S Meetei, J Das, C Roy Chaudhuri, and H Saha, "Quantitative recognition of flammable and toxic gases with artificial neural network using metal oxide gas sensors in embedded platform," *Eng. Sci. Technol. an Int. J.*, vol 18, no 2, pp 229–234, 2015

[6] W H Brattain and J Bardeen, "Surface properties of germanium," *Bell Syst. Tech. J.*, vol 32, no 1, pp 1–41, 1953

[7] T Seiyama, A Kato, K Fujiishi, and M Nagatani, "A new detector for gaseous components using semiconductive thin films ," *Anal. Chem.*, vol 34, no 11, pp 1502–1503, 1962

[8] C Wartelle, N Pereira Rodrigues, M Koudelka-Hep, and F Bedioui, "Amperometric fluidic microchip array sensing device for nitric oxide determination in solution," *Mater. Sci. Eng. C*, vol 26, no 2–3, pp 534–537, 2006

[9] S -M Park, S -L Zhang, and J -S Huh, "NO Sensing Characteristics of ZnO Nanorod Prepared by Ultrasound Radiation Method," *Korean J. Mater. Res.*, vol 18, no 7, pp 367–372, 2008

[10] Z Zhang, C Yin, L Yang, J Jiang, and Y Guo, "Optimizing the gas sensing characteristics of Co-doped SnO 2 thin film based hydrogen sensor," *J. Alloys Compd.*, vol 785, pp 819–825, 2019

[11] F Rathgeb and G Gauglitz, "Optical gas sensors in analytical chemistry: Applications, trends and general comments," *Encycl. Anal. Chem. Appl. Theory Instrum.*, 2006

[12] G Korotcenkov and B K Cho, "Metal oxide composites in conductometric gas sensors: Achievements and challenges," *Sensors Actuators, B Chem.*, vol 244, pp 182–210, 2017

[13] C Liu *et al.*, "A high-performance flexible gas sensor based on self-assembled PANI-CeO2 nanocomposite thin film for trace-level NH3 detection at room temperature," *Sensors Actuators, B Chem.*, vol 261, pp 587–597, 2018

[14] S Mahajan and S Jagtap, "Metal-oxide semiconductors for carbon monoxide (CO) gas sensing: A review," *Appl. Mater. Today*, vol 18, p 100483, 2020

[15] Y De Wang, X H Wu, Q Su, Y F Li, and Z L Zhou, "Ammonia-sensing characteristics of Pt and SiO2 doped SnO2 materials," *Solid. State. Electron.*, vol 45, no 2, pp 347–350, 2001

[16] S Ummartyotin and H Manuspiya, "A critical review on cellulose: From fundamental to an approach on sensor technology," *Renew. Sustain. Energy Rev.*, vol 41, pp 402–412, 2015

[17] M R Vilar *et al.*, "Development of nitric oxide sensor for asthma attack prevention," *Mater. Sci. Eng. C*, vol 26, no 2–3, pp 253–259, 2006

[18] H BISGAARD, L LOLAND, and J A N H ØJ, "NO in exhaled air of asthmatic children is reduced by the leukotriene receptor antagonist montelukast," *Am. J. Respir. Crit. Care Med.*, vol 160, no 4, pp 1227–1231, 1999

[19] A D Smith and D R Taylor, "Is exhaled nitric oxide measurement a useful clinical test in asthma?," *Curr. Opin. Allergy Clin. Immunol.*, vol 5, no 1, pp 49–56, 2005

[20] D R Taylor, M W Pijnenburg, A D Smith, and J C de Jongste, "Exhaled nitric oxide measurements: clinical application and interpretation," *Thorax*, vol 61, no 9, pp 817–827, 2006

[21] C Grimes *et al.*, "A sentinel sensor network for hydrogen sensing," *Sensors*, vol 3, no 3, pp 69–82, 2003

[22] M Z Jacobson, W G Colella, and D M Golden, "Cleaning the air and improving health with hydrogen fuel-cell vehicles," *Science (80-.).*, vol 308, no 5730, pp 1901–1905, 2005

[23] A M Bassam, A B Phillips, S R Turnock, and P A Wilson, "Development of a multi-scheme energy management strategy for a hybrid fuel cell driven passenger ship," *Int. J. Hydrogen Energy*, vol 42, no 1, pp 623–635, 2017

[24] L Boon-Brett *et al.*, "Identifying performance gaps in hydrogen safety sensor technology for automotive and stationary applications," *Int. J. Hydrogen Energy*, vol 35, no 1, pp 373–384, 2010

[25] W J Buttner, M B Post, R Burgess, and C Rivkin, "An overview of hydrogen safety sensors and requirements," *Int. J. Hydrogen Energy*, vol 36, no 3, pp 2462–2470, 2011

[26] K H Kim, E C Jeon, Y J Choi, and Y S Koo, "The emission characteristics and the related malodor intensities of gaseous reduced sulfur compounds (RSC) in a large industrial complex," *Atmos. Environ.*, vol 40, no 24, pp 4478–4490, 2006

[27] H Kimura, "Hydrogen sulfide: Its production and functions," *Exp. Physiol.*, vol 96, no 9, pp 833–835, 2011

[28] K H Kim, Y Choi, E Jeon, and Y Sunwoo, "Characterization of malodorous sulfur compounds in landfill gas," *Atmos. Environ.*, vol 39, no 6, pp 1103–1112, 2005

[29] M N Hughes, M N Centelles, and K P Moore, "Making and working with hydrogen sulfide The chemistry and generation of hydrogen sulfide in vitro and its measurement in vivo: A review," *Free Radic. Biol. Med.*, vol 47, no 10, pp 1346–1353, 2009

[30] S C K Misra, P Mathur, M Yadav, M K Tiwari, S C Garg, and P Tripathi, "Preparation and characterization of vacuum deposited semiconducting nanocrystalline polymeric thin film sensors for detection of HCl," *Polymer (Guildf).*, vol 45, no 25, pp 8623–8628, 2004

[31] C Yao *et al.*, "Sub-ppm CO detection in a sub-meter-long hollow-core negative curvature fiber using absorption spectroscopy at 2 3 μm," *Sensors Actuators, B Chem.*, vol 303, no October 2019, p 127238, 2020

[32] S H Nimkar, S P Agrawal, and S B Kondawar, "Fabrication of Electrospun Nanofibers of Titanium Dioxide Intercalated Polyaniline Nanocomposites for CO2 Gas Sensor," *Procedia Mater. Sci.*, vol 10, no Cnt 2014, pp 572–579, 2015

[33] M Imran, N Motta, and M Shafiei, "Electrospun one-dimensional nanostructures : a new horizon for gas sensing materials," no 2, 2018

[34] A Altindal, "ORGANlK YARI iLETKEN FiLMLERiN KARAKTERiZASYONU VE GAZ SENSORU OLARAK UYGULANMASI," 1999

[35] Z Li *et al.*, "Resistive-type hydrogen gas sensor based on TiO2: A review," *Int. J. Hydrogen Energy*, vol 43, no 45, pp 21114–21132, 2018

[36] "Wafer LiNbO3 Properties," 2020 [Online] Available: https://unitedcrystals com/LiNbO3Prop html [Accessed: 10-Jan-2020]

[37] "Gold Properties " [Online] Available: https://www lesker com/newweb/deposition_materials/depositionmaterials_sputtertargets_1 cfm?pgid=ti4 [Accessed: 24-Dec-2019]

[38] "Titanium Dioxide (TiO2) Sputtering Targets," 2020 [Online] Available: https://www lesker com/newweb/deposition_materials/depositionmaterials_sputtertargets_1 cfm?pgid=ti4#stq=&stp=1 [Accessed: 24-Oct-2019]

[39] "Zinc Oxide (ZnO) Sputtering Targets," 2020 [Online] Available: https://www lesker com/newweb/deposition_materials/depositionmaterials_sputtertargets_1 cfm?pgid=zn2 [Accessed: 24-Dec-2019]

[40] Comsol Multiphysics, "Heat Transfer, Comsol Multiphysics " 2020

[41] T K Roy, D Sanyal, D Bhowmick, and A Chakrabarti, "Temperature dependent resistivity study on zinc oxide and the role of defects," *Mater. Sci. Semicond. Process.*, vol 16, no 2, pp 332–336, 2013

[42] Y Zhang and J Han, "Microstructure and temperature coefficient of resistivity for ZnO ceramics doped with Al2O3," *Mater. Lett.*, vol 60, no 20, pp 2522–2525, 2006

A Low-cost Solution for Plasma Generation and Surface Treatment

A. OZCELIK[1] and M. K. KILINC[1]

[1] Aydın Adnan Menderes University, Aydın/Turkey, aozcelik@adu.edu.tr
[1] Aydın Adnan Menderes University, Aydın /Turkey, knklnc97@gmail.com

***Abstract* – In this work, practical and low-cost surface plasma treatment strategies are investigated. First, a consumer grade microwave is modified to obtain plasma in vacuum. Second, a high voltage generator is made into a hand-held device to form a portable and functional plasma generation source. The second approach is evaluated in terms of surface modification capability by treating different surfaces. Then, the surface treatment capacity is tested using contact angle measurements. It is found that, surface chemistry of various polymeric and ceramic surfaces can be changed based-on the changing contact angles which indicates a superhydrophobic to hydrophilic transition after plasma treatment.**

***Keywords* – Surface treatment, surface modification, PDMS bonding, contact angle, plasma generation**

I. INTRODUCTION

SURFACE treatment using plasma generation is important process for various applications in surface bonding, painting and chemical functionalization [1]–[3]. Especially in the field of microfluidics, polydimethylsiloxane (PDMS) is a commonly used material to fabricate microfluidic channels [4]–[6]. Bonding the PDMS layers to each other or to a glass slide is required to obtained a closed fluidic channel [7]–[10]. Plasma treatment of surfaces are usually carried out in plasma cleaners which consist of a closed chamber and a microwave generation mechanism [11], [12]. There are also hand-held high frequency plasma generation tools that can ignite and sustain a plasma in a small area in open air [12]. However, both of these approaches are costly and thus not accessible to small laboratories. Therefore, a simple and low cost solution is demanded for plasma generation and surface treatment in low-budget laboratories.

In general, a plasma is comprised of atoms, ions, electrons, ultraviolet light and higher energy atoms and molecules [13]. In a plasma, a very high amount of energy can exist. During the treatment of a surface with a plasma, all of these factors can interact with the surface. Plasma can be defined as the fourth state of matter which is formed by adding high energy to a gas.

In the event of a high voltage discharge in an air gap, free electrons, which are plentiful in the air, speeds up and by multiple collisions ionize the air. IF the electric discharge is powerful enough, collisions of rapidly moving electrons with atoms and molecules of the ambient gas yield a situation that momentum is conserved, and electron gathering occurs. When a plastic object is exposed in the discharge range, the electrons generated in the discharge bombard and affect the surface with energies 2 to 3 times higher than the required energy to break surface bonds of most materials. This creates highly reactive free chemical species. These chemical species in the presence of oxygen can react rapidly to form various chemical functional groups on the substrate surface. Functional groups occurring because of this oxidation reaction are very efficient at increasing surface energy and enhancing chemical bonding of surfaces. These chemical species include carbonyl (-C=O-). carboxyl (HOOC-), hydroperoxide (HOO-) and hydroxyl (HO-) groups [2], [3], [14].

In this work, we explore the low-cost and easy to use solutions for plasma generation. In this regard, we have designed and produced two plasma generation tools for surface treatment. In the first approach, we have modified a microwave oven to incorporate vacuum inside the microwave, and ignite plasma. In the second approach, we designed and fabricated a hand-held plasma generation tool to achieve surface treatment. We have used contact angle measurements on various materials after periodic plasma treatments using the hand-held plasma generation tool.

II. EXPERIMENTAL METHODS

In the first approach, a commercial microwave oven is modified to provide vacuum in a chamber inside. For this, a Teflon tube is inserted from the side airflow grid, and it is connected a home-made vacuum chamber inside the microwave. Any commercial microwave can be used in this experiment. For the vacuum generation, a standard vacuum pump is used which can provide a vacuum down to 13 Pa. For the vacuum chamber, an acrylic plate is used as the base to be connected to the vacuum line, and a glass beaker is placed on top of the plate. To provide a vacuum seal, a soft polymer layer made of polydimethylsiloxane (PDMS) is used between the glass beaker and the acrylic layer. PDMS is a soft polymer that provides a good seal under vacuum.

For the second approach, an arc generator coil with 1000kV capacity is used to induce a plasma in open air. This coil is placed inside a 3D printed holder to make it safer and easier to hold. A fused deposition modelling (FDM) type 3D printer (i3 Mega, Anycubic, Shenzhen, China) is used for 3D printing the plastic parts of the system. The following printing parameters are set in the slicing software; print speed: 50 mm per second, infill: 100% with zig zag pattern, cooling: 100%, retraction: enabled, layer height: 0.1 mm, and wall line count: 3. Standard 1.75 mm diameter polylactic acid (PLA) filament is used with 210 °C at the print-head and 60 °C at the print bed. An adjustable DC power supply is used to provide 7 V, 0.5 Amps to the arc

generating coil. This hand-held plasma generator is then evaluated by treating surfaces of different materials. Glass, PDMS, PVC, acrylic and polystyrene are used for the surface treatment experiments. The surfaces are treated with 0.5 seconds intervals. Surface changes are measured using contact angle measurements. For this, 5 microliter water droplets are dispensed on the untreated and treated surfaces, and a side-view picture is taken using an iPhone 7 smartphone. Then, ImageJ is used to measure the contact angles of the water droplets.

III. RESULTS AND DISCUSSIONS

In the microwave generated plasma experiments, we encountered several problems. The main problem was the compatibility of the materials that we could place inside the microwave oven. That is, any material such as adaptors, base plate, vacuum chamber and tubing were required to be unaffected by the microwave. Otherwise, the materials were heated up and damaged inside the microwave. For example, Teflon was compatible with microwave and was unaffected for any duration of microwave exposure. However, some of the plastics including O-rings for sealing functions were damaged and sometimes burnt inside the microwave. After finding the suitable materials and building the vacuum chamber inside the microwave, we tested the system with the microwave.

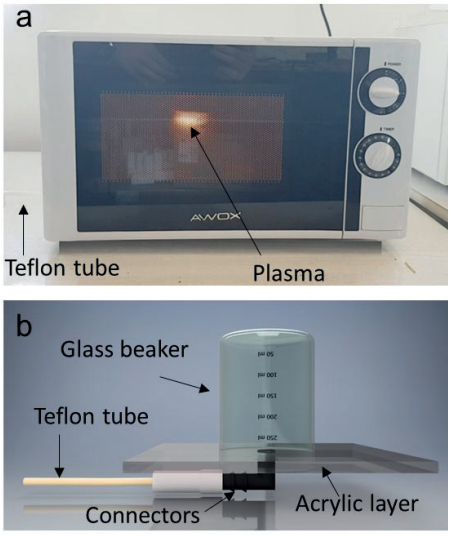


Figure 1: A modified microwave is used to generate vacuum plasma. a) A plasma is generated after applying vacuum and microwaves. b) A custom-made vacuum chamber is used to generate the required vacuum to ignite the plasma.

We found out that the level of the base vacuum was critical for igniting the plasma. This is understandable from the basics of a plasma. If there are excessive amount of air inside the vacuum chamber, the mean free path of the ion and electrons will be much smaller, and the plasma will not be sustained due to the insufficient amount of continuous ionization of the atoms and molecules inside the chamber. The vacuum pump applied in this experiment is rated for a minimum 13 Pa vacuum level. However, it took about 5 to 10 minutes to reach to this level of vacuum. In addition, considering the non-ideal conditions and potential small leaks, the vacuum was not optimum inside the chamber. Nevertheless, after pumping the vacuum chamber down for about 10 minutes, we obtained plasma that could be sustained as long as the microwave was turned on as shown in Figure 1. Overall, we have found this approach cumbersome for quick and practical surface treatment requirements. Therefore, we looked for a more practical solution using a hand-held approach.

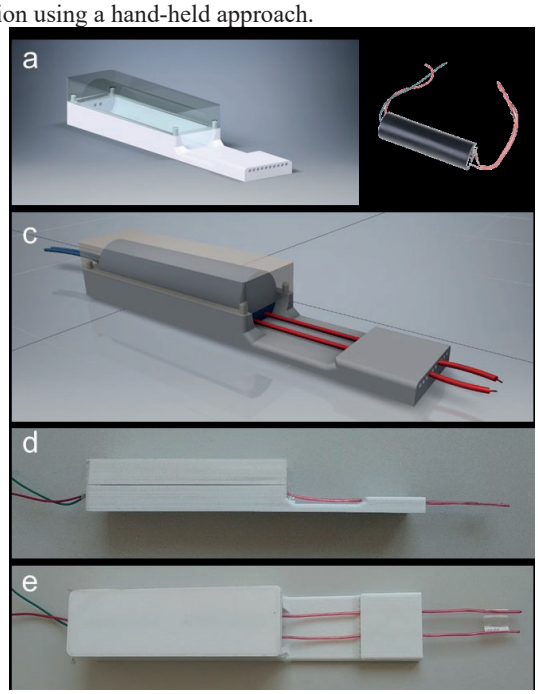


Figure 2: a) Design schematics, b) ignition coil, c) assembled device schematic, d) side view of the actual device and e) top view of the actual device.

For the portable plasma generation approach, a high frequency and high voltage ignition coil (1000 kV) is placed inside the 3D printed case as shown in Figure 2. This device can be run via a lithium ion battery to become a truly portable device. However, we simulated the battery with a variable DC power generator because of the unavailability of a suitable lithium ion battery with specific voltage requirements for the coil. The holder shown in Figure 2 was designed to have various holes at the tip for adjusting the spacing between the high voltage discharge output leads. For the ideal case, we set the distance to 10 mm which are taped together to provide further rigidity. In Figure 3, the plasma generator was turned off and on. A plasma is ignited and sustained at the tip of the output leads.

After successfully generating a viable tool for generating plasma, we decided to apply it on various surfaces to investigate time dependent effect of plasma treatment. We chose readily available materials in our lab for the surface treatment experiments which are glass slides, poly acrylic, polyvinyl chloride (PVC), polydimethylsiloxane (PDMS), and polystyrene. These materials are frequently used in industrial applications. For example, PDMS is a rapid prototyping material for fabricating microfluidic channels using soft lithography and replica molding methods. PDMS and glass are finally bond together to obtain a sealed microfluidic channel. Plasma treatment in this process is critical to activate both PDMS and glass slide surfaces.

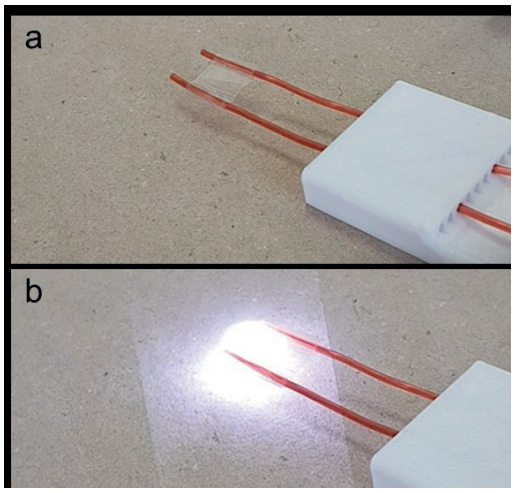


Figure 3: The hand-held device is a) turned off and b) turned on. A glare is visible in b) due to the refreshment rate of the applied voltage

In the surface treatment experiments, we exposed the surfaces to the generated plasma with 0.5 second intervals. For surface evaluation method, we used 5 microliter droplets and measured the contact angles of the droplets. An example measurement of the contact angle is shown in Figure 4. Before the surface plasma treatment, the 5 microliter droplet does not wet the surface. After the plasma exposure to the glass surface, droplet wets the surface, and spread by the reduction of the surface energy. Overall, the surface treatment made the surface more hydrophilic. This transformation is due to the change in the surface radicals by the high energy plasma.

After testing the capabilities of the hand-held plasma generation device, we tried to expose and modify surfaces of 5 different materials. Namely, glass, acrylic, PDMS, polystyrene, and PVC were chosen due their common use in various applications. Before applying the plasma, all the surfaces were initially evaluated by the contact angle measurements. As seen in Figure 5, untreated surfaces demonstrated a hydrophobic nature by not wetting the 5 microliter droplet on the surface. Then, these surfaces were treated with the plasma for 0.5 seconds. The contact angle measurements after the first 0.5 second plasma treatment yielded much more wetting hydrophilic surfaces. The following 0.5 second plasma treatment made the surfaces even more hydrophilic. This treatment was applied 3 times on the surfaces, and the contact angles were measured after each treatment.

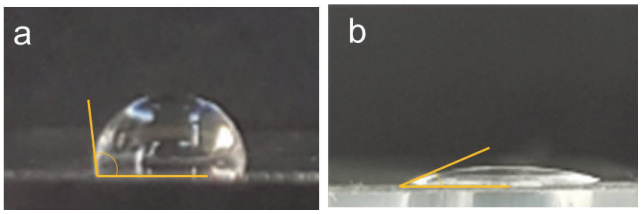


Figure 4: Demonstration of contact angle measurement a) before and b) after plasma treatment on a glass surface.

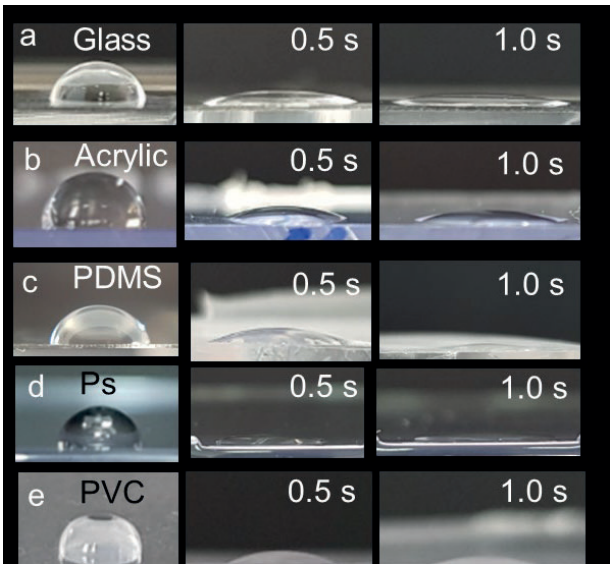


Figure 5: Contact angle data for the a) glass, b) acrylic, c) PDMS, d) polystyrene, and e) PVC. Initially the surfaces were hydrophobic, and after 0.5 and 1 second of plasma treatment, they became hydrophilic.

Characterization of the contacts angles measurements for the 45 different materials is given in Figure 6. Here, each measurement on a certain material was repeated at least 3 times. The average value of these 3 measurements were used in the graph in Figure 6. The error bars were calculated from the standard deviation of the repeated measurements. As it is observed from the Figure 6, after the first 0.5 second plasma treatment, contact angles dropped significantly which was corresponding to a more hydrophilic surface. The second and third plasma treatments changed the contact angle less prominently. Eventually, after the overall plasma treatment, surfaces became completely hydrophilic. The highest initial contact angle was measured for the PDMS, and the lowest contact angle after the complete plasma treatment was recorded for the polystyrene surface.

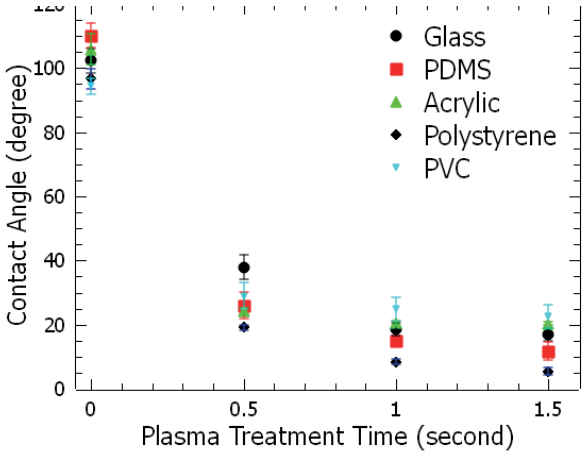


Figure 6: Characterization of the contact angles measurments of the different surfaces.Standard deviation of at least 3 measruements for each data point was used as the error bars.

IV. CONCLUSIONS

In conclusion, we investigated the potential solutions for a low-cost plasma generation tool. The microwave approach was more difficult to apply in terms of practicality due to the incompatibility of various materials in microwave. Then, we devised a hand-held instrument to generate plasma. This form of plasma generation was applied to treat surfaces of glass slides, poly acrylic, polyvinyl chloride (PVC), polydimethylsiloxane (PDMS), and polystyrene. The contact angle measurements revealed the effective modification of the surfaces using the simple plasma generation tool. This hand-held plasma source cost less than 10 USD in fabrication which is orders of magnitude lower in cost compared to the commercial hand-held plasma generation tools. Therefore, this approach can be adopted by the low-resource setting laboratories that are in need of a practical way of applying plasma to surfaces. The current device can generate plasma in a 1 centimeter square area at once. It is also possible to further develop the method for treating larger areas in shorter times. Overall, this is a low-cost and practical solution for surface plasma applications.

ACKNOWLEDGMENT

This work was supported by The Scientific and Technological Research Council of Turkey (TÜBİTAK) 3501 grant (Project No. 218M257)

REFERENCES

[1] R L Puurunen, "Surface chemistry of atomic layer deposition: A case study for the trimethylaluminum/water process," *J. Appl. Phys.*, vol 97, no 12, p 121301, 2005, doi: 10 1063/1 1940727

[2] P Chu, "Plasma-surface modification of biomaterials," *Mater. Sci. Eng. R Reports*, vol 36, no 5–6, pp 143–206, Mar 2002, doi: 10 1016/S0927-796X(02)00004-9

[3] E M Liston, L Martinu, and M R Wertheimer, "Plasma surface modification of polymers for improved adhesion: a critical review," *J. Adhes. Sci. Technol.*, vol 7, no 10, pp 1091–1127, Jan 1993, doi: 10 1163/156856193X00600

[4] Y Xia and G M Whitesides, "SOFT LITHOGRAPHY," *Annu. Rev. Mater. Sci.*, vol 28, no 1, pp 153–184, Aug 1998, doi: 10 1146/annurev matsci 28 1 153

[5] D Qin, Y Xia, and G M Whitesides, "Soft lithography for micro- and nanoscale patterning ," *Nat. Protoc.*, vol 5, no 3, pp 491–502, Mar 2010, doi: 10 1038/nprot 2009 234

[6] P Zuo, X Li, D C Dominguez, and B C Ye, "A PDMS/paper/glass hybrid microfluidic biochip integrated with aptamer-functionalized graphene oxide nano-biosensors for one-step multiplexed pathogen detection," *Lab Chip*, vol 13, no 19, pp 3921–3928, 2013, doi: 10 1039/c3lc50654a

[7] D J Laser and J G Santiago, "A review of micropumps," *J. Micromechanics Microengineering*, vol 14, no 6, 2004, doi: 10 1088/0960-1317/14/6/R01

[8] N Skov and H Bruus, "Modeling of Microdevices for SAW-Based Acoustophoresis — A Study of Boundary Conditions," *Micromachines*, vol 7, no 10, p 182, Oct 2016, doi: 10 3390/mi7100182

[9] D Kilinc, A Blasiak, and G U Lee, "Microtechnologies for studying the role of mechanics in axon growth and guidance," *Front. Cell. Neurosci.*, vol 9, no JULY, pp 1–8, 2015, doi: 10 3389/fncel 2015 00282

[10] C -Y Lee, C -L Chang, Y -N Wang, and L -M Fu, "Microfluidic Mixing: A Review," *Int. J. Mol. Sci.*, vol 12, no 12, pp 3263–3287, May 2011, doi: 10 3390/ijms12053263

[11] J Musschoot *et al.*, "Comparison of Thermal and Plasma-Enhanced ALD/CVD of Vanadium Pentoxide," *J. Electrochem. Soc.*, vol 156, no 7, p P122, 2009, doi: 10 1149/1 3133169

[12] H Conrads and M Schmidt, "Plasma generation and plasma sources," *Plasma Sources Sci. Technol.*, vol 9, no 4, pp 441–454, Nov 2000, doi: 10 1088/0963-0252/9/4/301

[13] S O Macheret, M N Shneider, and R B Miles, "Modeling of air plasma generation by repetitive high-voltage nanosecond pulses," *IEEE Trans. Plasma Sci.*, vol 30, no 3, pp 1301–1314, Jun 2002, doi: 10 1109/TPS 2002 802142

[14] S Guruvenket, G M Rao, M Komath, and A M Raichur, "Plasma surface modification of polystyrene and polyethylene," *Appl. Surf. Sci.*, vol 236, no 1–4, pp 278–284, Sep 2004, doi: 10 1016/j apsusc 2004 04 033

Numerical Study and Modelling of Turbulent Flow Over Backward-Facing Steps

T. KARASU[1]

[1]Eskişehir Osmangazi University, Meşelik, 26480 Eskişehir/Turkey, tkarasu@ogu.edu.tr

Abstract - **This original research paper presents the results of an extensive investigation of numerical study and modelling of steady, incompressible, two-dimensional, separating and reattaching turbulent flow over backward-facing steps at three different Reynolds numbers. Employing the finite-volume method with a hybrid scheme, a computer program based on the SIMPLE (Semi-Implicit Method for Pressure Linked Equations) algorithm has been developed. Numerical solution of conservation equations of mass and momentum, together with the standard k-ε turbulence model, are obtained using an iterative numerical solution technique. Near the solid boundaries, wall-functions are employed. Numerical predictions for local streamwise velocity, turbulence kinetic energy, turbulence kinetic energy dissipation rate, effective viscosity profiles, locus of flow reversal, bottom wall static-pressure coefficient, wall-shear stress and friction coefficient distributions along top and bottom walls of the backward-facing step flow configurations are presented and compared with experimental measurements. The results of numerical study are generally in good agreement with experimental data.**

Keywords - **Turbulent flow, k-ε turbulence model, Predictions.**

I. Introduction

THE separating and reattaching turbulent recirculating flow in a two-dimensional backward-facing step flow configuration shown in Figure 1, is of great practical importance. It is an excellent case for analysing the basic physical phenomena of separation and reattachment because it is the simplest such flow. As revealed in Figure 1, three fundamental flow regimes are present: a reattaching shear-layer, a recirculating region, and a stagnation flow region. These regimes rigorously test the validity of turbulence models, which are, for the most part, tuned to specific homogeneous shear flows. Turbulent flow in backward-facing steps, sudden expansions and pipes has been studied both experimentally and theoretically by numerous investigators, among the most interesting ones are the works of [1-44]. In the present numerical study, the experimental measurements of [1,3,11] have been employed for comparison with the numerical predictions. The main objective of this original investigation is to obtain numerical predictions for steady, incompressible, two-dimensional turbulent recirculating flow over backward-facing steps using the standard k-ε turbulence model [45] with a wall-functions boundary condition, and to compare the results of numerical computations with available experimental measurements in the literature for validation.

II. Mathematical and Physical Model

A. Governing Equations and Turbulence Model

With reference to Figure 1, the mathematical and physical model employed in the numerical study of steady, incompressible, and two-dimensional separating and reattaching turbulent flow over backward-facing steps, requires the simultaneous solution of the governing equations, together with the turbulence model equations. The transport equations representing the conservation of mass, momentum, turbulence kinetic energy and its dissipation rate in steady state cartesian coordinate system are cast into the following general form:

$$\frac{\partial}{\partial x}\left(\rho u \phi\right)+\frac{\partial}{\partial y}\left(\rho v \phi\right)-\frac{\partial}{\partial x}\left(\Gamma_\phi \frac{\partial \phi}{\partial x}\right)-\frac{\partial}{\partial y}\left(\Gamma_\phi \frac{\partial \phi}{\partial y}\right)=S_\phi \quad (1)$$

This equation represents a transport equation for a general variable ϕ. The variables u and v are the local time-averaged velocity components in the horizontal (x) and normal (y) directions, respectively, ρ is the fluid density and Γ_ϕ is the transport coefficient. The final term, S_ϕ , is the source term. The variables, ϕ , necessary in this numerical study are u, v, k and ε. Here, k and ε stand for the kinetic energy of turbulence and its dissipation rate, respectively. The turbulence model employed in the present study is the k-ε model of [45]. If ϕ is set equal to unity and Γ_ϕ and S_ϕ to zero, equation (1) reduces to the equation of continuity. Pressure is derived from the pressure correction equation [46,47]. The transport equations, coefficients and the source terms are summarised in Table 1.

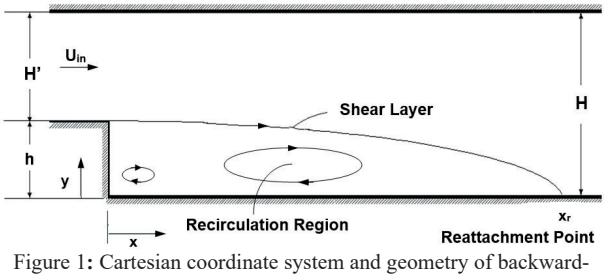


Figure 1: Cartesian coordinate system and geometry of backward-facing step flow channel.

Table 1. Transport equations, coefficients and source terms

Variable ϕ	Γ_ϕ	S_ϕ
1	0	0
u	$\mu_e = \mu + \mu_t$	$-\frac{\partial P}{\partial x} + \frac{\partial}{\partial x}\left(\mu_e \frac{\partial u}{\partial x}\right) + \frac{\partial}{\partial y}\left(\mu_e \frac{\partial v}{\partial x}\right)$
v	μ_e	$-\frac{\partial P}{\partial y} + \frac{\partial}{\partial x}\left(\mu_e \frac{\partial u}{\partial y}\right) + \frac{\partial}{\partial y}\left(\mu_e \frac{\partial v}{\partial y}\right)$
k	$\frac{\mu_e}{\sigma_k}$	$G - \rho\varepsilon$
ε	$\frac{\mu_e}{\sigma_\varepsilon}$	$\frac{\varepsilon}{k}(C_1 G - C_2 \rho\varepsilon)$

Note : 1. $\mu_t = C_\mu \rho k^2/\varepsilon$

2. Turbulence model constants are assigned the following values (Launder and Spalding, 1974):
$C_\mu = 0.09, C_1 = 1.44, C_2 = 1.92, \sigma_k = 1.0, \sigma_\varepsilon = 1.3$

3. $G = \mu_t \left\{ 2\left[\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial v}{\partial y}\right)^2 \right] + \left(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x}\right)^2 \right\}$

B. Boundary Conditions

With reference to Figure 1, the boundary conditions for two-dimensional separating and reattaching turbulent flow over a backward-facing step studied are given below. Three different boundary conditions were used**:** inlet plane, outlet plane, step, bottom and top solid walls. At the step inlet plane either streamwise velocity distribution or a uniform velocity distribution from the experimental measurements is specified, while the velocity in the normal direction is set equal to zero. Empirical relationships are employed to assign uniform entrance values to turbulence quantities k and ε**;** that is, $k=(0.002\text{-}0.005)\, u_b^2$ and $\varepsilon = (C_\mu k^{3/2} / 0.03H)$, where u_b is the bulk velocity at inlet and H is the channel height. The outlet plane was located sufficiently far away from the recirculation region so that fully-developed conditions could be assumed to prevail at the outlet**;** that is, all the streamwise gradients of the dependent variables vanish at the outlet plane and the normal velocity is assumed zero. At the step, bottom and top solid walls, velocity components u, v and turbulence quantities k and ε are set to zero. The values of k and ε at the near-wall grid points are evaluated using the wall-functions of [45]. Initial field values throughout the computational domain were specified properly so as not to cause numerical divergence.

C. *Numerical Solution Procedure*

In this numerical study, employing the finite-volume approach, a computer program based on the SIMPLE algorithm of [46] has been developed. The partial differential equations (1) are discretised by a control-volume-based finite-difference method with a hybrid scheme. The finite-volume forms of the partial differential equations coupled with the boundary conditions are solved iteratively using a line-by-line solution procedure in conjunction with a tridiagonal matrix form [48].

D. Computational Details

The numerical computations were carried out on a Pentium 4 CPU 1**.**60 GHz personal computer. The computational grid distributions for all of the backward-facing step flows studied are depicted in Figures 2, 13 and 24. All the computational grids employed were non-uniformly distributed with dense grid-line concentrations in the recirculation region and near the walls of the backward-facing step flow channel. Staggered control volumes were used for streamwise and normal velocity components. All other quantities of interest were calculated at the grid points. The line-by-line method was used to obtain converged solutions iteratively. Underrelaxation factors were employed to procure numerical stability with values of 0**.**3,0**.**3,0**.**8,0**.**8,0**.**5 and 0**.**3 for u, v, k, ε, P and μ_e, respectively. The convergence criterion adopted in the present computations was that the summation of the absolute values of the mass residual in the entire computational domain be less than a prescribed value of 10^{-5}. Grid tests were performed with different grid sizes to obtain an optimum grid-independent solution for each flow case studied. All the computations presented in this numerical study are grid-independent. Table 2 summarises details of computational requirements for all flow cases investigated. In this table, Re is the Reynolds number and N is the number of iterations performed to obtain a converged solution.

Table 2**:** Flow case, Reynolds number, grid size and number of iterations.

Flow case of	Re	Grid Size (x) × (y)	N
Stevenson et. al. [3]	172 518	40 × 30	567
Adams and Eaton [1]	36 000	40 × 30	274
Ötügen [11]	16 600	40 × 30	423

III. PRESENTATION AND DISCUSSION OF RESULTS

Numerical predictions have been conducted for two-dimensional separating and reattaching turbulent flow over backward-facing steps at three different Reynolds numbers, and the results of numerical predictions were compared with the experimental measurements of [1], [3], and [11]. With air used as the working fluid in the backward-facing step flow channel, the experimental measurements of [3] are first selected as the comparison basis for this numerical study. The backward-facing step flow configuration has an expansion ratio of H / H**'** = 2. The Reynolds number of the flow at the inlet is Re = 172 518 ($Re = u_{in}\, h / \nu$, where u_{in} is the average free stream velocity at the step inlet and h is the step height). For this flow situation, the computational domain is extended to a streamwise distance of x=20h downstream from the inlet plane of the backward-facing step flow channel. The numerical grid distribution for the backward-facing step flow geometry of [3] is displayed in Figure 2. The predicted local streamwise velocity profiles along the backward-facing step flow channel in dimensionless form u/u_{in}, and normal distance y/h, at streamwise locations ranging from x/h=0**.**5 to 20 across the channel have been compared with the experimental measurements of [3] in Figure 3. The local streamwise velocity is made dimensionless with respect to reference average free stream velocity u_{in} at the step inlet. Examination of Figure 3 indicates that the predicted non-dimensional

streamwise velocity profiles exhibit reasonable agreement with the corresponding experimental ones. In addition, this figure also shows how the predicted local streamwise velocity profiles develop along the backward-facing step flow channel. Figures 4, 5 and 6 reveal the predicted profiles of turbulence kinetic energy, turbulence kinetic energy dissipation rate and effective viscosity, respectively, along the backward-facing step flow channel of [3] at the same downstream locations as in Figure 3. Here, the predicted profiles are normalised, respectively, with respect to square of average free stream velocity u_{in}^2 at the step inlet, maximum values of turbulence kinetic energy dissipation rate ε_{max} and effective viscosity $\mu_{e\,max}$ in the flow field. These figures show how the predicted turbulence kinetic energy, dissipation rate of turbulence kinetic energy and effective viscosity profiles develop along the channel. The predicted locus of flow reversal (u=0) for the backward-facing step flow channel is presented in Figure 7. The normal distance y measured from the bottom wall is normalised to step height (h) and plotted as a function of downstream distance normalised to step height (x/h). In this figure, the streamwise velocity (u) is zero on the circles, positive above the circles and negative below the circles. As can be seen from the figure, the predicted flow reattachment length is found to occur at about x_r/h=5.75 downstream from the inlet plane, whereas the experimental flow reattachment length is estimated to occur at about x_r/h=7. The source of this problem is a limitation of the standard k-ε turbulence model. The model does not account for the time lapse between extra turbulent energy being supplied and the effect being felt in the dissipating motions. The predicted wall-shear stress distribution along the backward-facing step flow channel bottom-wall in terms of τ_{wb}/τ_{wdb}, as a function of downstream distance x/h, is depicted in Figure 8. Here, the wall-shear stress has been made dimensionless with respect to its value at the outlet plane (τ_{wdb}). As can be seen from the figure, starting from the inlet plane of the channel until at about x/h=5.75 the wall-shear stress is negative. This is due to the negative streamwise velocities (-u) in the separated-reattaching flow region.

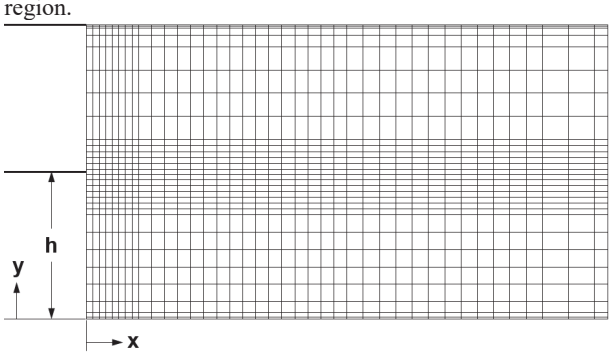


Figure 2: Numerical grid distribution for backward-facing step flow geometry of [3].

The bottom wall-shear stress, as it takes negative values, increases rapidly until it attains its maximum negative value, and then decreases sharply to its zero value at about x/h=5.75, afterwhich it takes positive values and increases continuously.

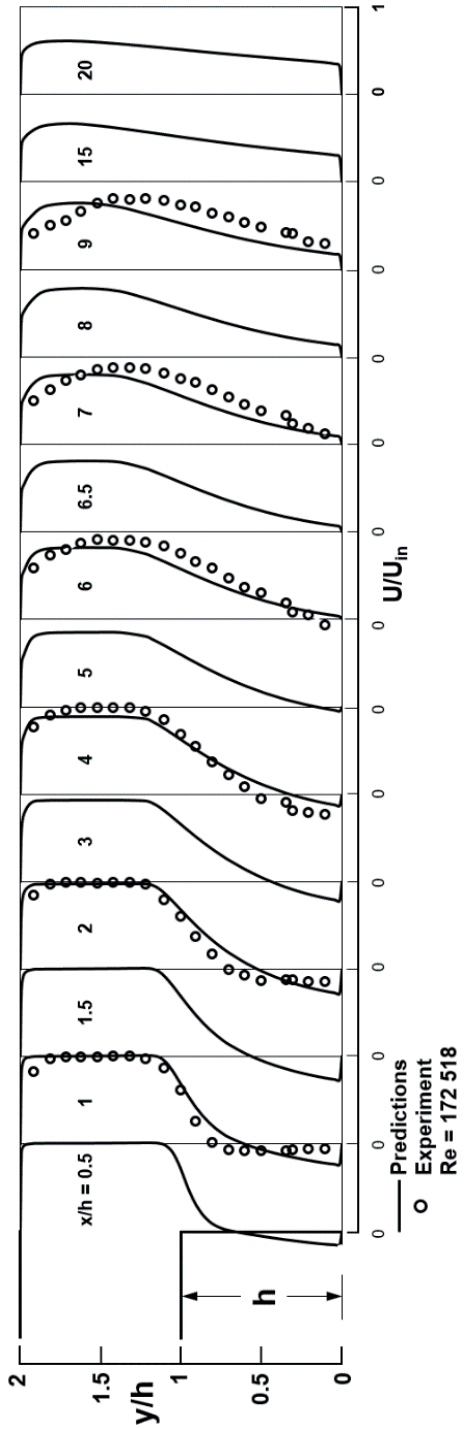


Figure 3: Comparison of predicted dimensionless streamwise velocity profiles along backward-facing step flow channel with experimental measurements of [3].

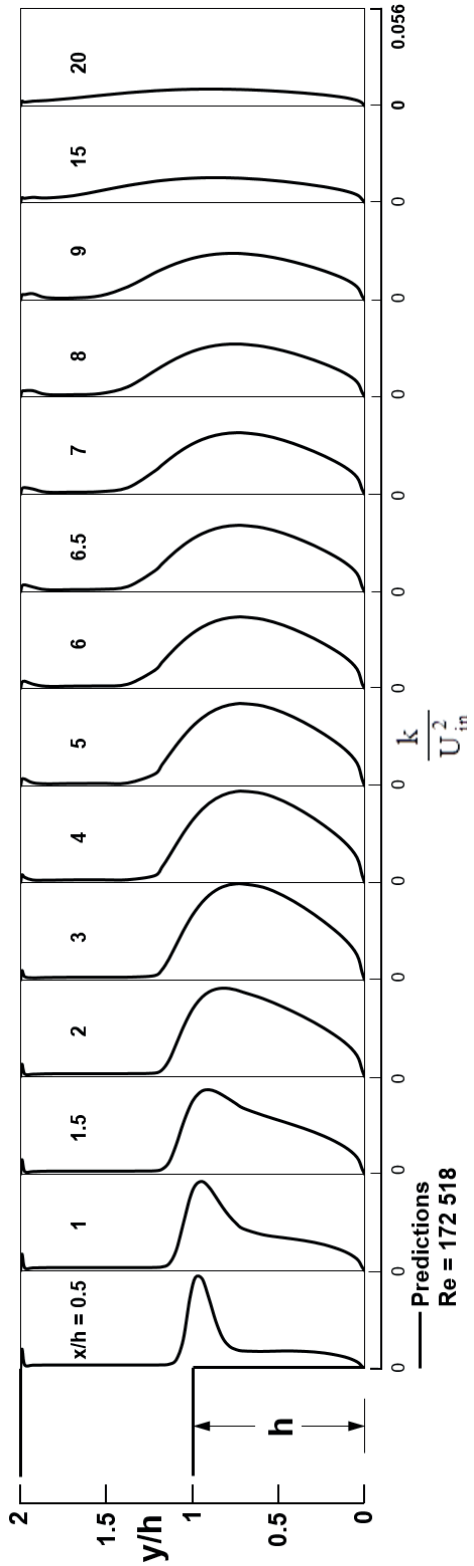


Figure 4: Predicted profiles of normalised turbulence kinetic energy along backward-facing step flow channel of [3].

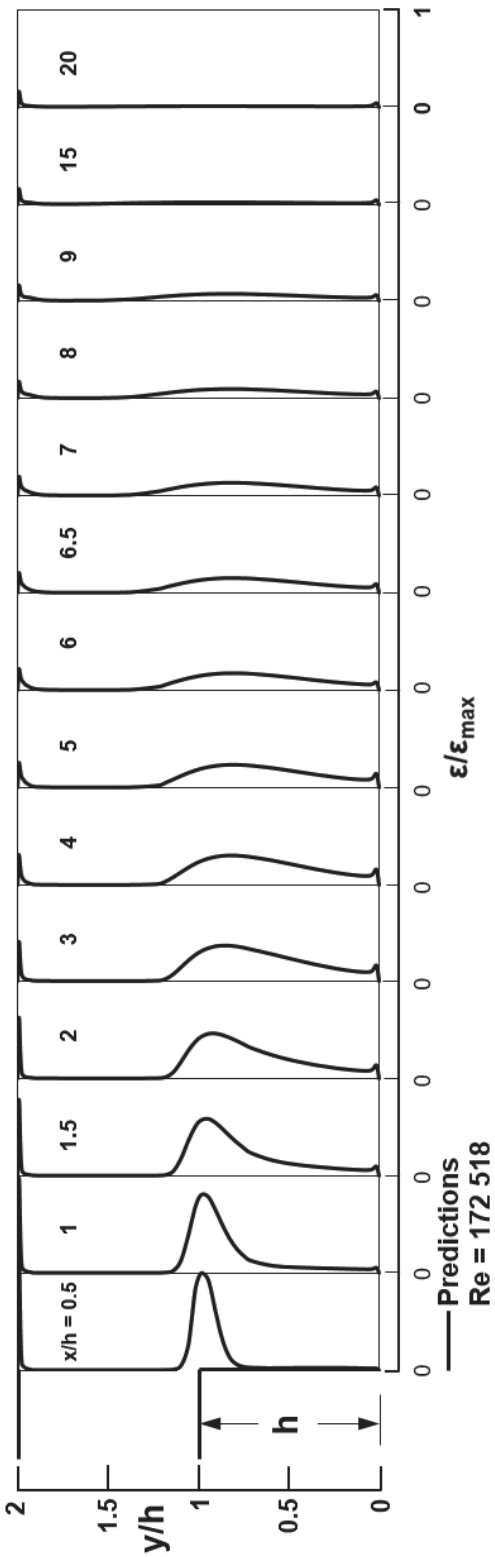


Figure 5: Predicted profiles of normalised turbulence kinetic energy dissipation rate along backward-facing step flow channel of [3].

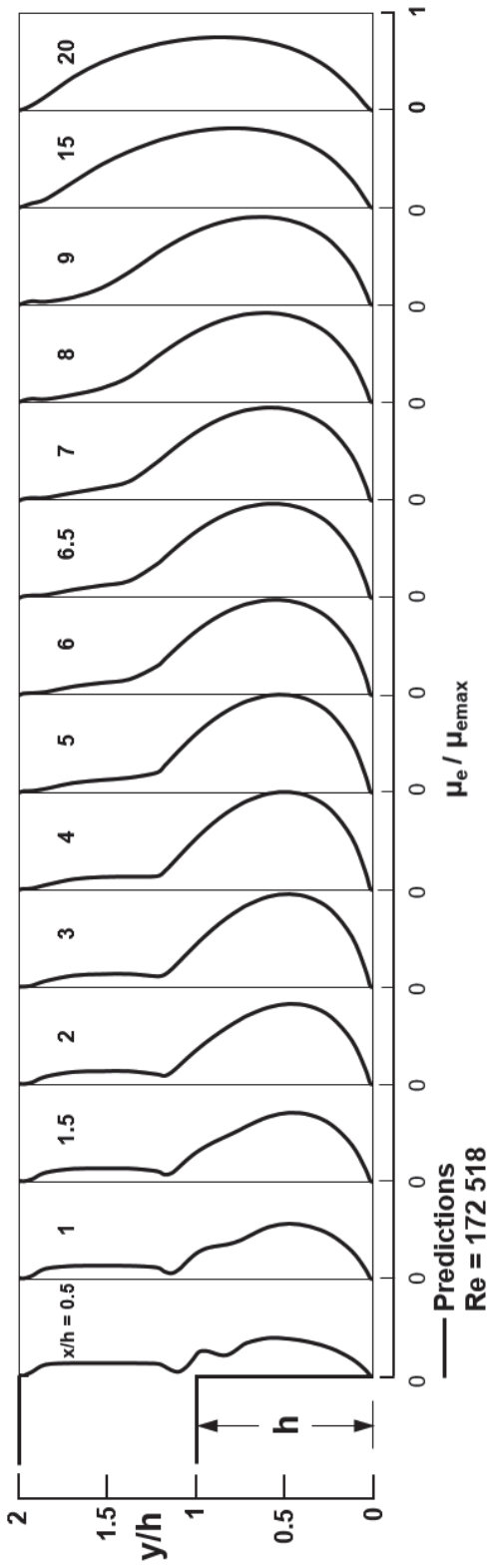


Figure 6: Predicted profiles of normalised effective viscosity along backward-facing step flow channel of [3].

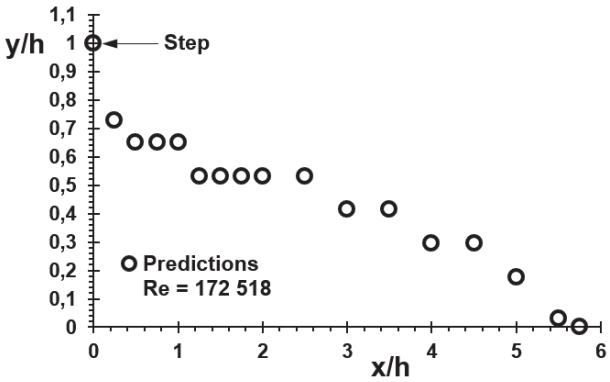


Figure 7: Predicted locus of flow reversal (u=0) for backward-facing step flow of [3].

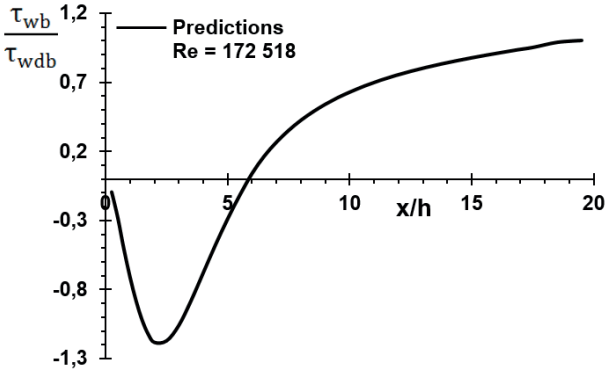


Figure 8: Predicted distribution of dimensionless wall-shear stress along bottom-wall of backward-facing step flow channel of [3].

The predicted distribution of friction coefficient (C_{fb} = 2 τ_{wb} / ρ u_{in}^2) along the channel bottom-wall, as a function of downstream distance x/h, is plotted in Figure 9. As seen from the figure, the friction coefficient exhibits the same trend as the wall-shear stress shown in Figure 8. The predicted variation of wall static-pressure coefficient [C_p = (P - P_{ref}) / 0.5 ρ u_{in}^2] along the channel bottom-wall, as a function of downstream distance x/h, is given in Figure 10. As seen from the figure, the bottom-wall static-pressure coefficient takes negative values just from the inlet plane of the channel until its zero value at about x/h=2.25, afterwhich it takes positive values and increases continuously up to downstream location x/h=19.5. In Figure 11, the predicted distribution of wall-shear stress along the channel top-wall in dimensionless form of τ_{wt} / τ_{wdt}, as a function of downstream distance x/h, is displayed. Here, the wall-shear stress has been normalised with respect to its value at the outlet plane (τ_{wdt}). As can be seen from the figure, starting from the inlet plane of the channel until at x/h=19.5 the wall-shear stress decreases continuously. Finally, the computed variation of friction coefficient (C_{ft} = 2 τ_{wt} / ρ u_{in}^2) along the channel top-wall, as a function of downstream distance x/h, is presented in Figure 12. As seen from the figure, the friction coefficient reveals the same trend as the wall-shear stress shown in Figure 11.

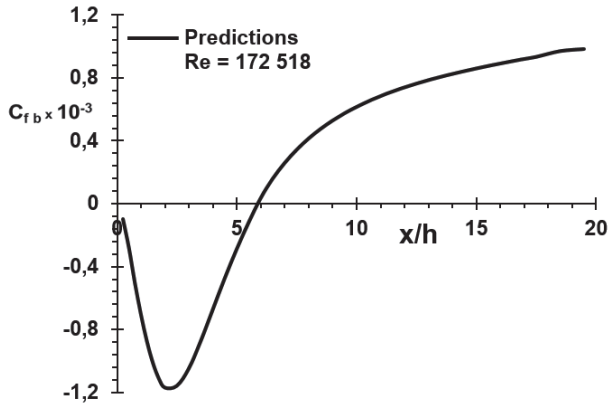


Figure 9: Predicted distribution of friction coefficient along bottom-wall of backward-facing step flow channel of [3].

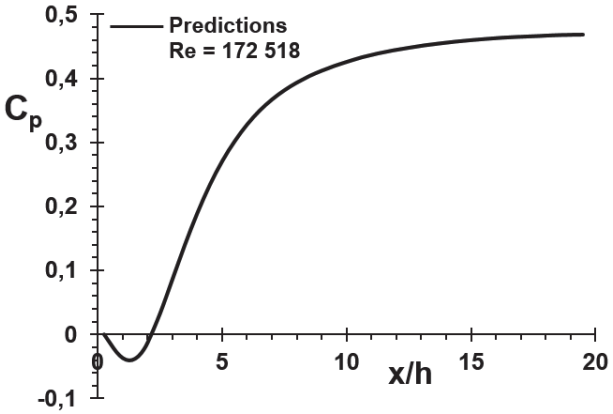


Figure 10: Predicted variation of wall static-pressure coefficient (C_p) along bottom-wall of backward-facing step flow channel of [3].

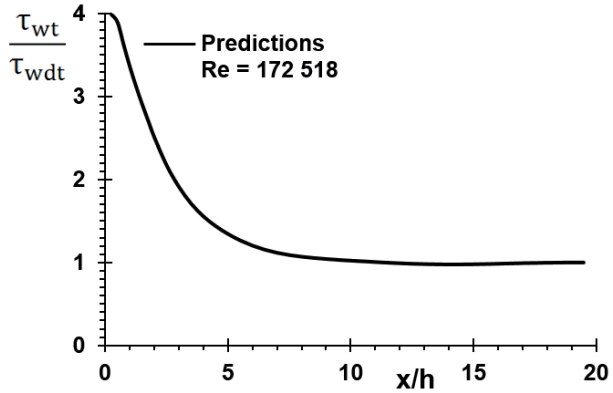


Figure 11: Predicted distribution of dimensionless wall-shear stress along top-wall of backward-facing step flow channel of [3].

The second flow configuration corresponds to the two-dimensional backward-facing step flow experimental study of [1] with air used as the working fluid. The backward-facing step flow channel has an expansion ratio of H / H' = 1.25. The Reynolds number of the flow is Re = 36 000 (Re = u_{ref} h / ν, where u_{ref} is the upstream reference velocity measured at 12.7 cm upstream of the step and h is the step height).

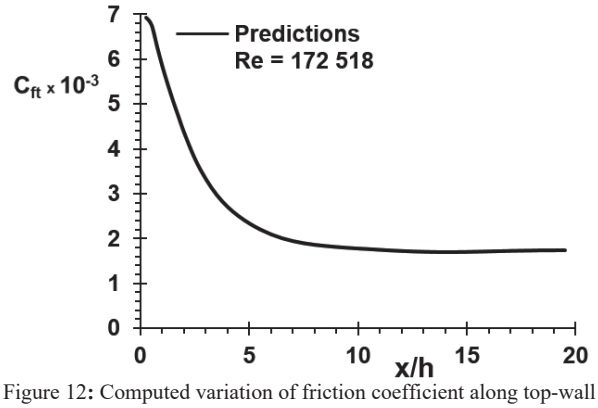


Figure 12: Computed variation of friction coefficient along top-wall of backward-facing step flow channel of [3].

For this flow case, the calculation domain is extended to a streamwise distance of x=20h downstream from the inlet plane of the backward-facing step flow channel. The computational grid distribution for the backward-facing step flow of [1] is depicted in Figure 13. The predicted local streamwise velocity profiles for the step flow in terms of u/u_{ref}, and transverse distance y/h, at streamwise positions ranging from x/h=0.16 to 18 across the channel are presented in Figure 14 and compared with experimental data of [1]. Here, the local streamwise velocity is normalised with respect to upstream reference velocity u_{ref}, measured at 12.7 cm upstream of the step in the freestream. As seen from the figure, the predicted streamwise velocity profiles are in pretty good agreement with the corresponding experimental ones. In addition, this figure also reveals how the predicted streamwise velocity profiles develop along the backward-facing step flow channel. Figures 15, 16 and 17 display the predicted profiles of turbulence kinetic energy, turbulence kinetic energy dissipation rate and effective viscosity, respectively, along the step flow channel of [1] at the same sections as in Figure 14. Here, the predicted profiles are normalised, respectively, with respect to square of upstream reference velocity u_{ref}^2, maximum values of turbulence kinetic energy dissipation rate ε_{max} and effective viscosity $\mu_{e\,max}$ in the flow field. These figures indicate how the predicted turbulence kinetic energy, dissipation rate of turbulence kinetic energy and effective viscosity profiles develop along the channel.

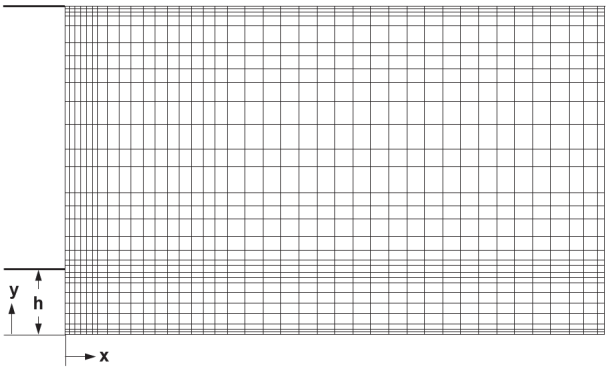


Figure 13: Computational grid distribution for backward-facing step flow geometry of [1].

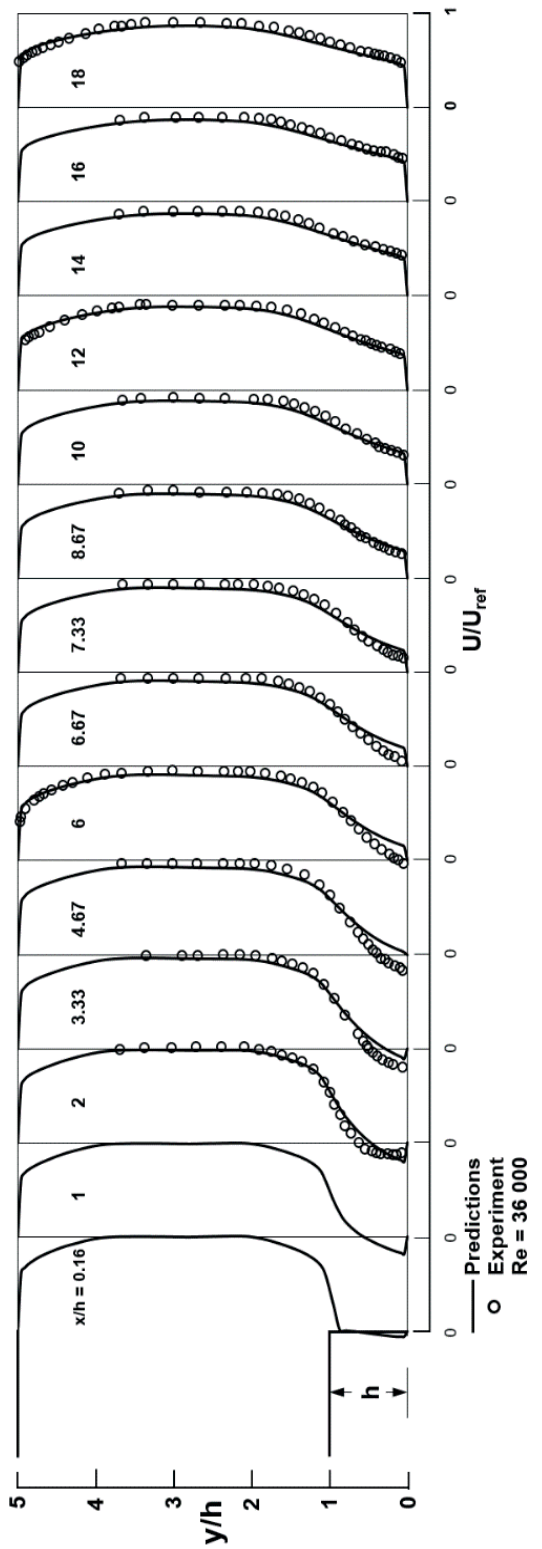


Figure 14: Comparison of predicted dimensionless streamwise velocity profiles along backward-facing step flow channel with experimental data of [1].

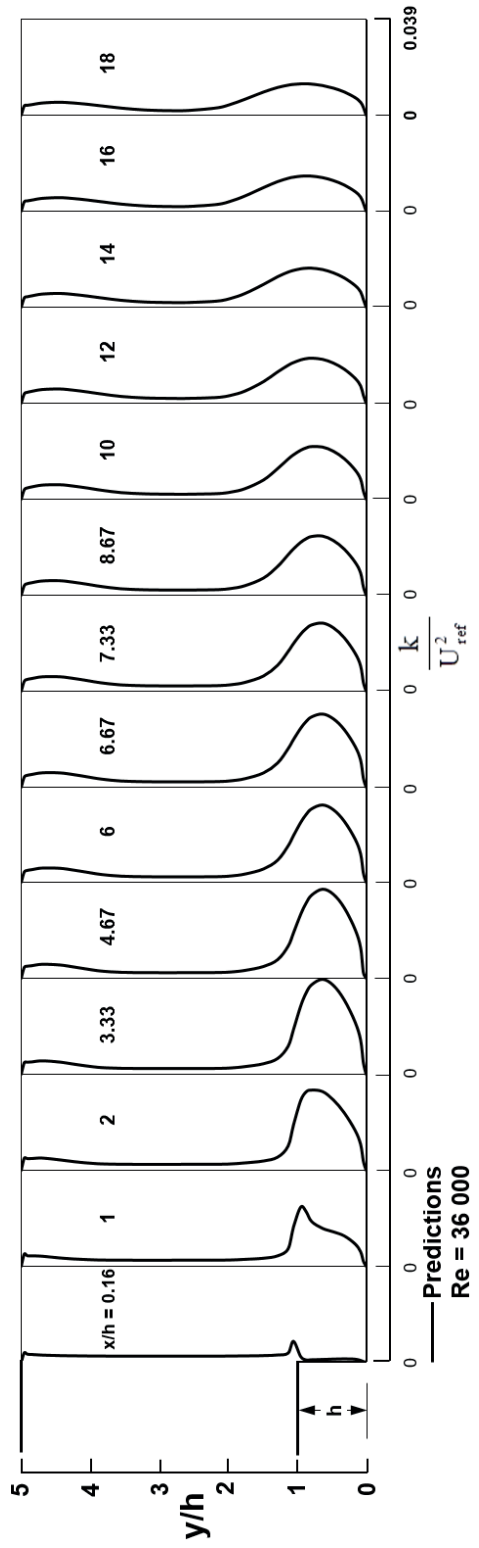


Figure 15: Predicted profiles of normalised turbulence kinetic energy along backward-facing step flow channel of [1].

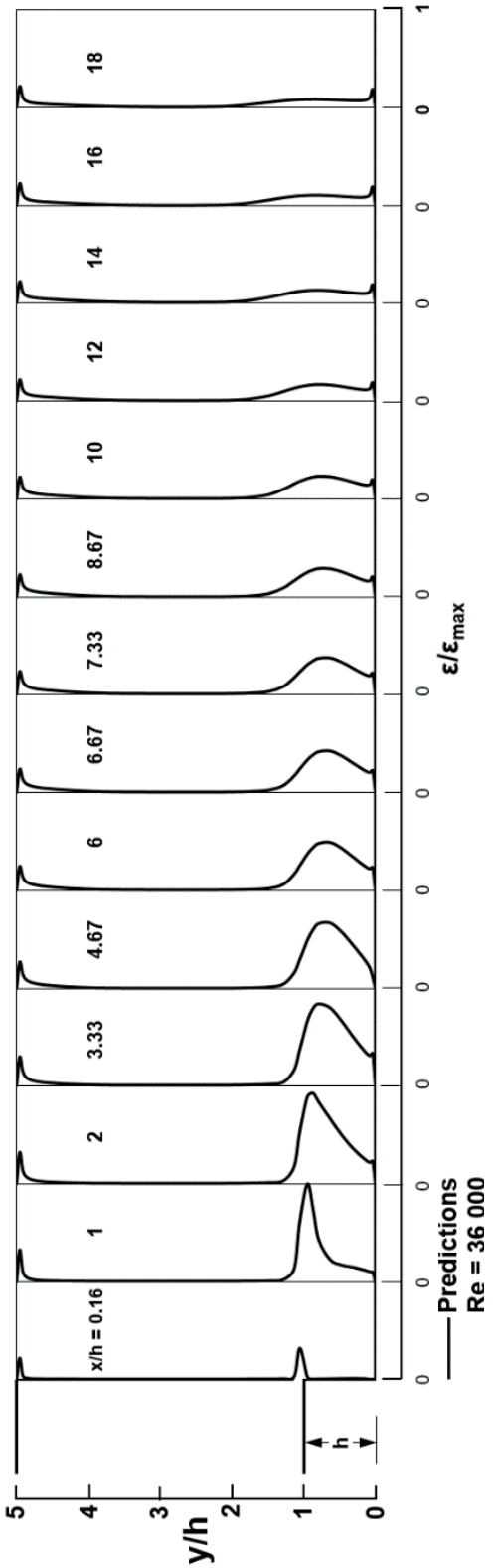


Figure 16: Predicted profiles of normalised turbulence kinetic energy dissipation rate along backward-facing step flow channel of [1].

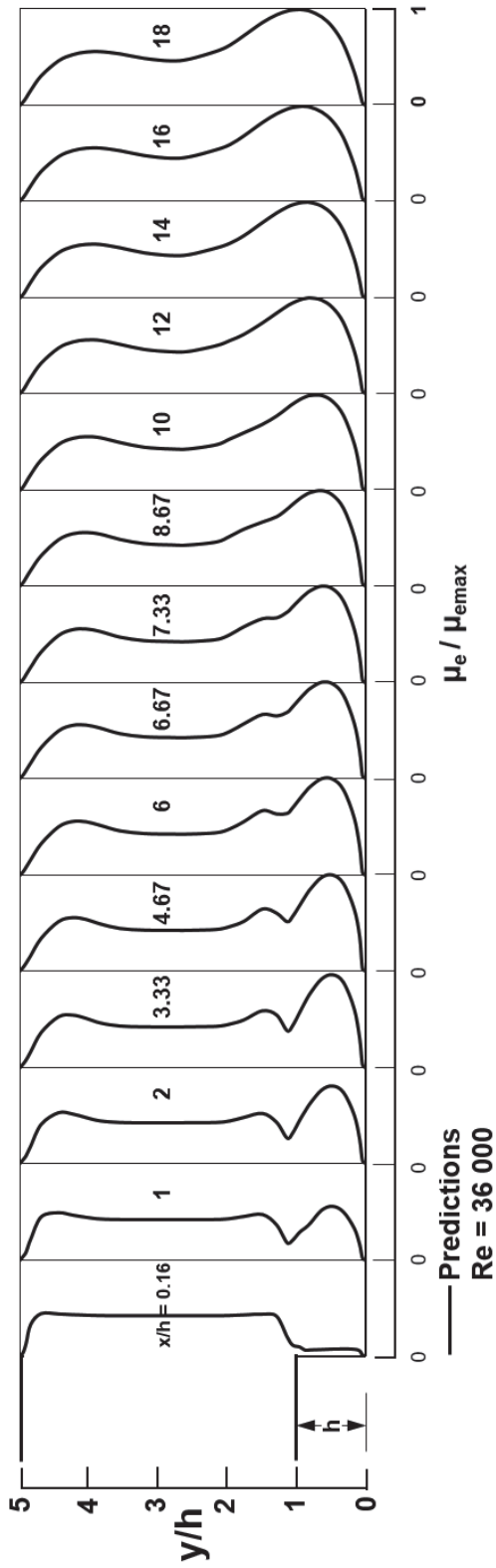


Figure 17: Predicted profiles of normalised effective viscosity along backward-facing step flow channel of [1].

The predicted locus of flow reversal (u=0) for the backward-facing step flow is given in Figure 18. In this figure, the streamwise velocity (u) is zero on the circles, positive above the circles and negative below the circles. For this flow case, as seen from the figure, the predicted flow reattachment length is found to occur at about x_r/h=4.446 downstream from the inlet plane, while the experimental flow reattachment length is reported to occur at x_r/h=6.6. The source of this problem is a limitation of the standard k-ε turbulence model. The model does not account for the time lapse between extra turbulent energy being supplied and the effect being felt in the dissipating motions. The computed wall-shear stress distribution along the backward-facing step flow channel bottom-wall in dimensionless form of τ_{wb}/τ_{wdb}, as a function of downstream distance x/h, is manifested in Figure 19. Here, the wall-shear stress is nondimensionalised with respect to its value at the outlet plane (τ_{wdb}). As seen from the figure, starting from the inlet plane of the channel until at about x/h=4.6 the wall-shear stress is negative. This is due to the negative streamwise velocities (-u) in the recirculation zone. The bottom wall-shear stress, as it takes negative values, increases rapidly until it reaches its maximum negative value, and then decreases sharply to its zero value at about x/h=4.6, afterwhich it takes positive values and increases continuously. The computed distribution of friction coefficient ($C_{fb} = 2\ \tau_{wb} / \rho\ u_{ref}^2$) along the channel bottom-wall, as a function of downstream distance x/h, is compared with the experimental measurements of [1] in Figure 20. As seen from the figure, the friction coefficient indicates the same trend as the experimental measurements of [1] and shows reasonable agreement. The predicted variation of wall static-pressure coefficient [$C_p = (P - P_{ref}) / 0.5\ \rho\ u_{ref}^2$] along the channel bottom-wall, as a function of downstream distance x/h, is plotted in Figure 21. As seen from the figure, the bottom-wall static-pressure coefficient takes negative values just from the inlet plane of the channel until its zero value at about x/h=2, afterwhich it takes positive values and increases continuously up to downstream location x/h=19.7. In Figure 22, the predicted distribution of wall-shear stress along the channel top-wall in terms of τ_{wt} / τ_{wdt}, as a function of downstream distance x/h, is presented. Here, the wall-shear stress has been normalised with respect to its value at the outlet plane (τ_{wdt}). As can be seen from the figure, starting from the inlet plane of the channel until at about x/h=1 the wall-shear stress increases sharply until it reaches its maximum positive value, and thereafter decreases up to x/h=12, afterwhich it remains nearly constant. Finally, the computed variation of friction coefficient ($C_{ft} = 2\ \tau_{wt} / \rho\ u_{ref}^2$) along the channel top-wall, as a function of downstream distance x/h, is depicted in Figure 23. As seen from the figure, the friction coefficient exhibits the same trend as the wall-shear stress indicated in Figure 22.

With air used as the working fluid in the backward-facing step flow channel, the experimental study for turbulent flow conducted by [11] is thirdly selected as the test case for this computational study. The backward-facing step flow channel has an expansion ratio of H / H' = 2. The Reynolds number of the flow at the inlet is Re = 16 600 ($Re = u_o H' / \nu$, where u_o is the reference free stream velocity and H' is the upstream channel height). For this flow situation, the computational domain is extended to a streamwise distance of x=25h downstream from the inlet plane of the backward-facing step flow channel. The numerical grid distribution for the backward-facing step flow of [11] is presented in Figure 24.

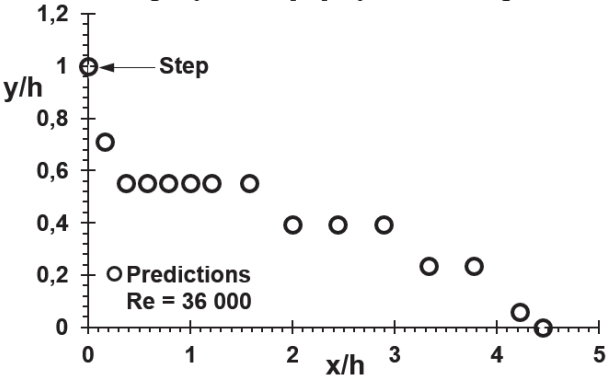


Figure 18: Predicted locus of flow reversal (u=0) for backward-facing step flow of [1].

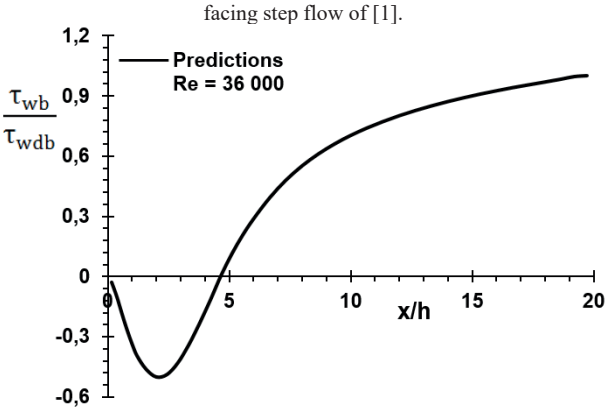


Figure 19: Computed distribution of dimensionless wall-shear stress along bottom-wall of backward-facing step flow channel of [1].

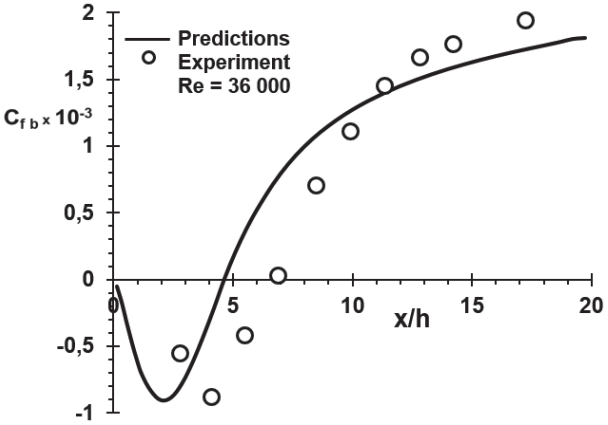


Figure 20: Comparison of computed distribution of friction coefficient along bottom-wall of backward-facing step flow channel with experimental measurements of [1].

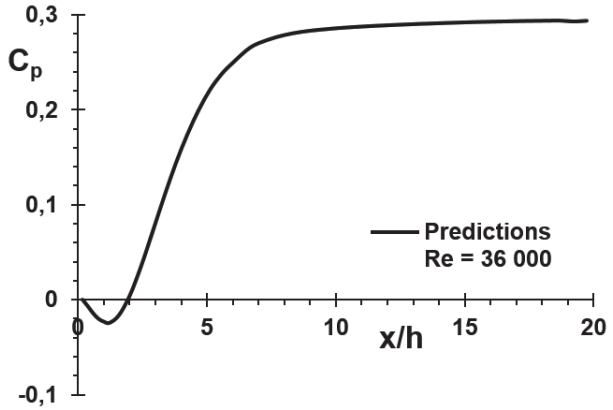


Figure 21: Predicted variation of wall static-pressure coefficient (C_p) along bottom-wall of backward-facing step flow channel of [1].

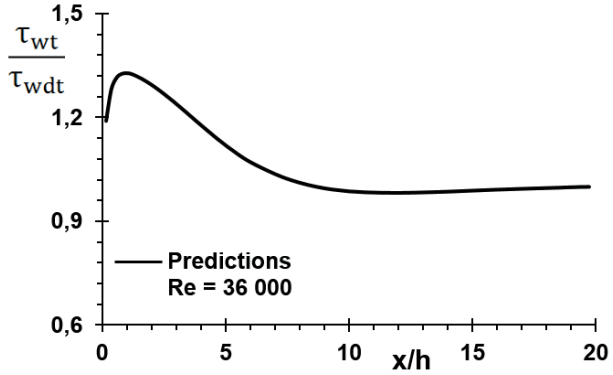


Figure 22: Predicted distribution of dimensionless wall-shear stress along top-wall of backward-facing step flow channel of [1].

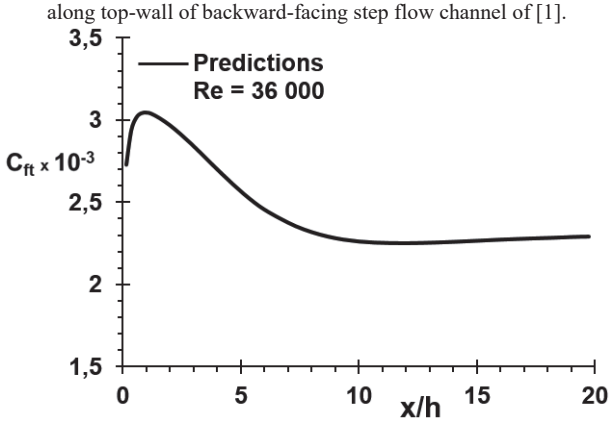


Figure 23: Computed variation of friction coefficient along top-wall of backward-facing step flow channel of [1].

The predicted local streamwise velocity profiles along the backward-facing step flow channel in dimensionless form u/u_m, and normal distance y/h, at streamwise locations ranging from x/h=0.3 to 25 across the channel have been compared with the experimental measurements of [11] in Figures 25, 26, 27 and 28. Here, the local streamwise velocity is made dimensionless with respect to maximum velocity u_m in profile at each cross-section. As seen from the figures, the predicted non-dimensional streamwise velocity profiles are in reasonable accord with the corresponding experimental ones taken in the separated shear layer only, and as the flow develops along the channel the agreement becomes better. Figures 29, 30 and 31 exhibit the predicted profiles of turbulence kinetic energy, dissipation rate of turbulence kinetic energy and effective viscosity, respectively, along the step flow channel of [11] at the same downstream locations as in Figure 25. Here, the predicted profiles are normalised, respectively, with respect to square of reference free stream velocity u_o^2, maximum values of turbulence kinetic energy dissipation rate ε_{max} and effective viscosity $\mu_{e\ max}$ in the flow field. These figures denote how the predicted profiles develop along the channel. The predicted locus of flow reversal (u=0) for the step flow is indicated in Figure 32. In this figure, the streamwise velocity (u) is zero on the circles, positive above the circles and negative below the circles. For this flow situation, the predicted flow reattachment length is found to occur at x_r/h=5 downstream from the inlet plane of the channel. The predicted wall-shear stress distribution along the step flow channel bottom-wall in terms of τ_{wb}/τ_{wdb}, as a function of downstream distance x/h, is shown in Figure 33. Here, the wall-shear stress has been made dimensionless with respect to its value at the outlet plane (τ_{wdb}). As seen from the figure, starting from the inlet plane of the channel until at about x/h=5.3 the wall-shear stress is negative. This is due to the negative streamwise velocities (-u) in the recirculation zone. The bottom wall-shear stress, as it takes negative values, increases rapidly until it attains its maximum negative value, and then decreases sharply to its zero value at about x/h=5.3, afterwhich it takes positive values and increases continuously. The predicted distribution of friction coefficient ($C_{fb} = 2\ \tau_{wb} / \rho\ u_o^2$) along the channel bottom-wall, as a function of downstream distance x/h, is plotted in Figure 34. As seen from the figure, the friction coefficient reveals the same trend as the wall-shear stress shown in Figure 33. The predicted distribution of wall static-pressure coefficient [$C_p = (P - P_{ref}) / 0.5\ \rho\ u_o^2$] along the channel bottom-wall, as a function of downstream distance x/h, is displayed in Figure 35. As seen from the figure, the bottom-wall static-pressure coefficient takes negative values just from the inlet plane of the channel until its zero value at x/h=2, afterwhich it takes positive values and increases continuously up to x/h=17.5. Beyond this position, it remains constant.

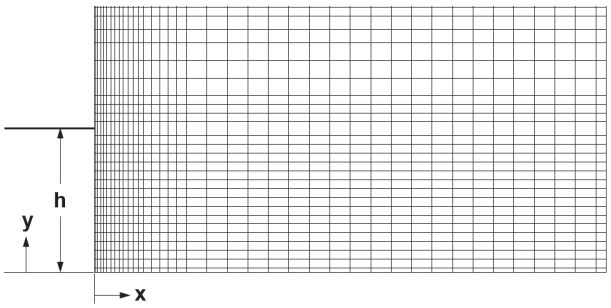


Figure 24: Numerical grid distribution for backward-facing step flow geometry of [11].

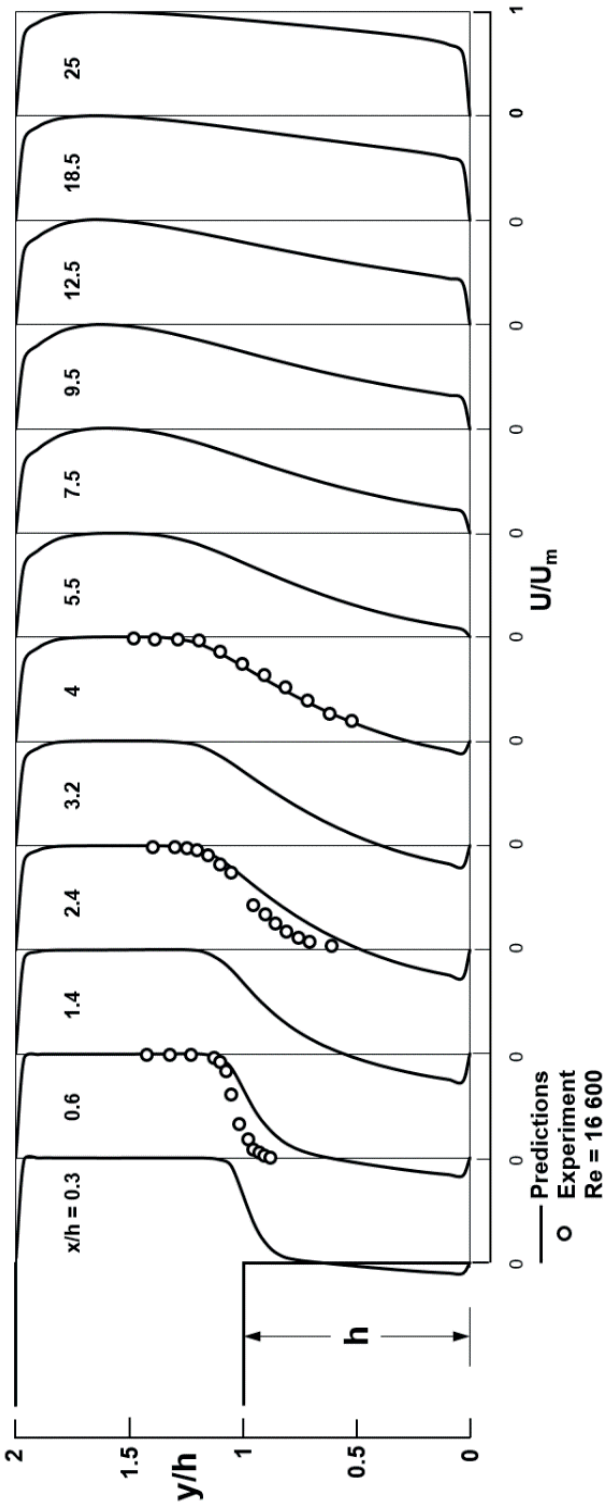


Figure 25: Comparison of predicted dimensionless streamwise velocity profiles along backward-facing step flow channel with experimental measurements of [11].

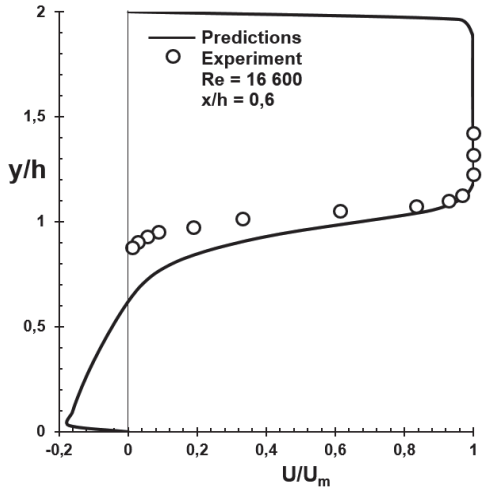


Figure 26: Comparison of predicted dimensionless streamwise velocity profile at x/h=0.6 with experimental measurements of [11].

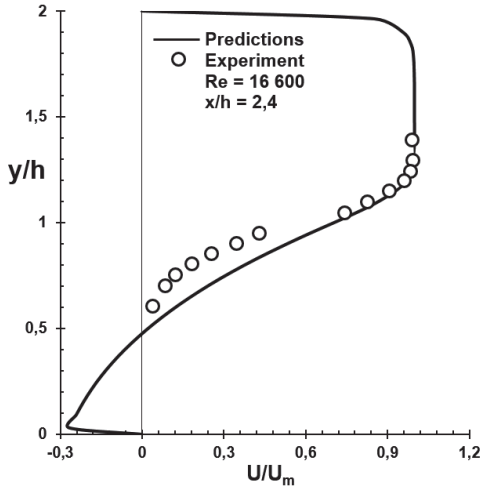


Figure 27: Comparison of predicted dimensionless streamwise velocity profile at x/h=2.4 with experimental measurements of [11].

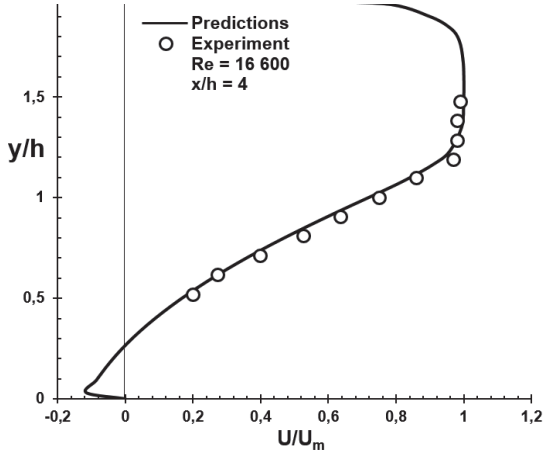


Figure 28: Comparison of predicted dimensionless streamwise velocity profile at x/h=4 with experimental measurements of [11].

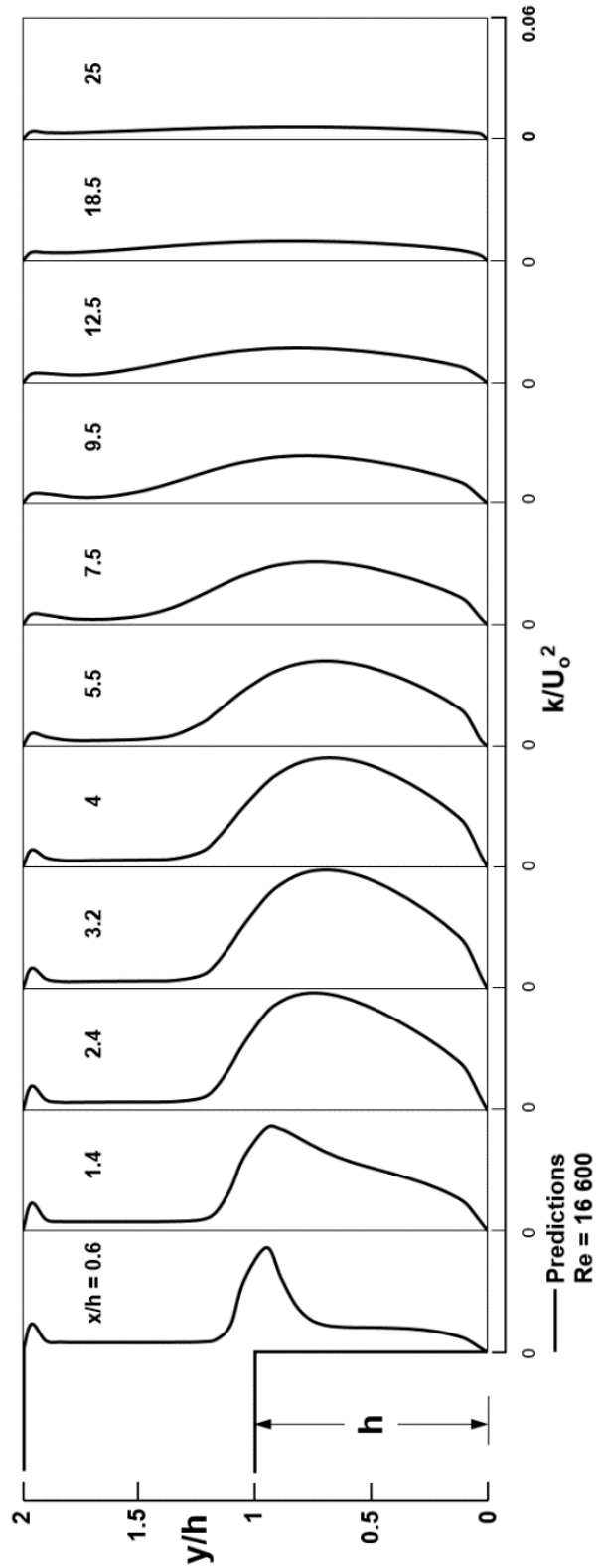


Figure 29**:** Predicted profiles of normalised turbulence kinetic energy along backward-facing step flow channel of [11].

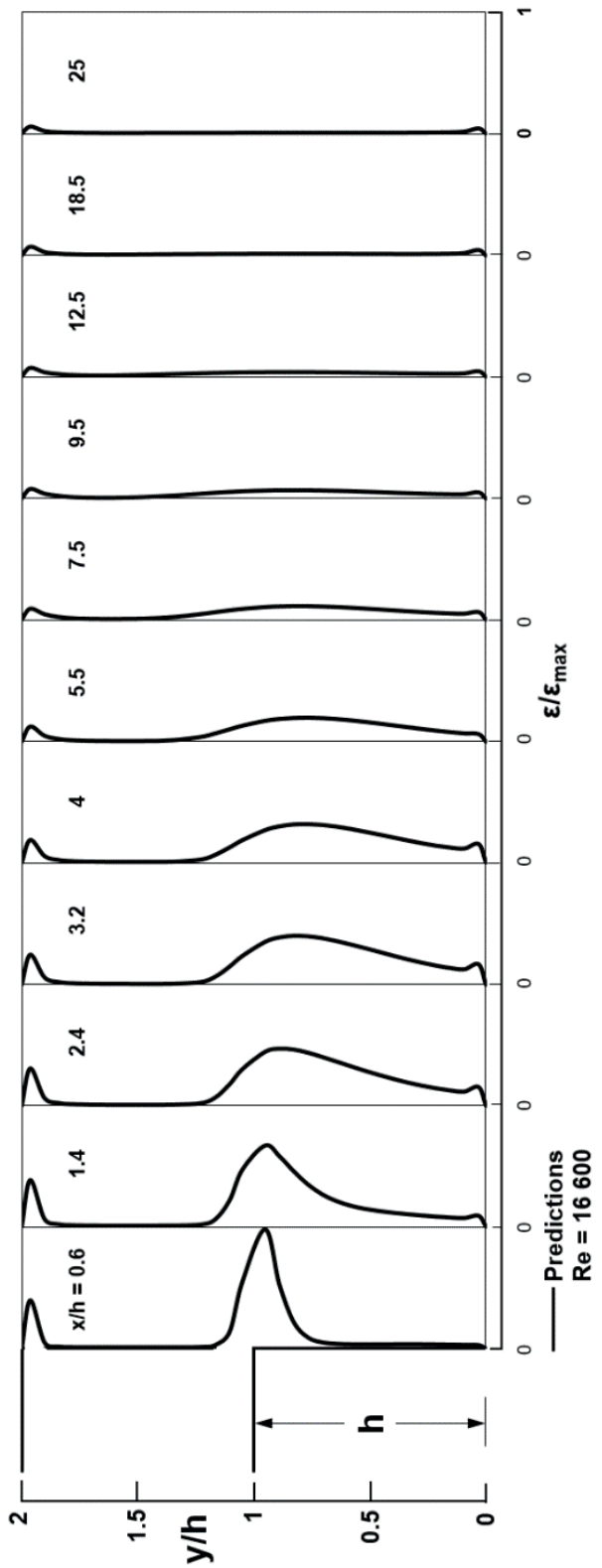


Figure 30**:** Predicted profiles of normalised turbulence kinetic energy dissipation rate along backward-facing step flow channel of [11].

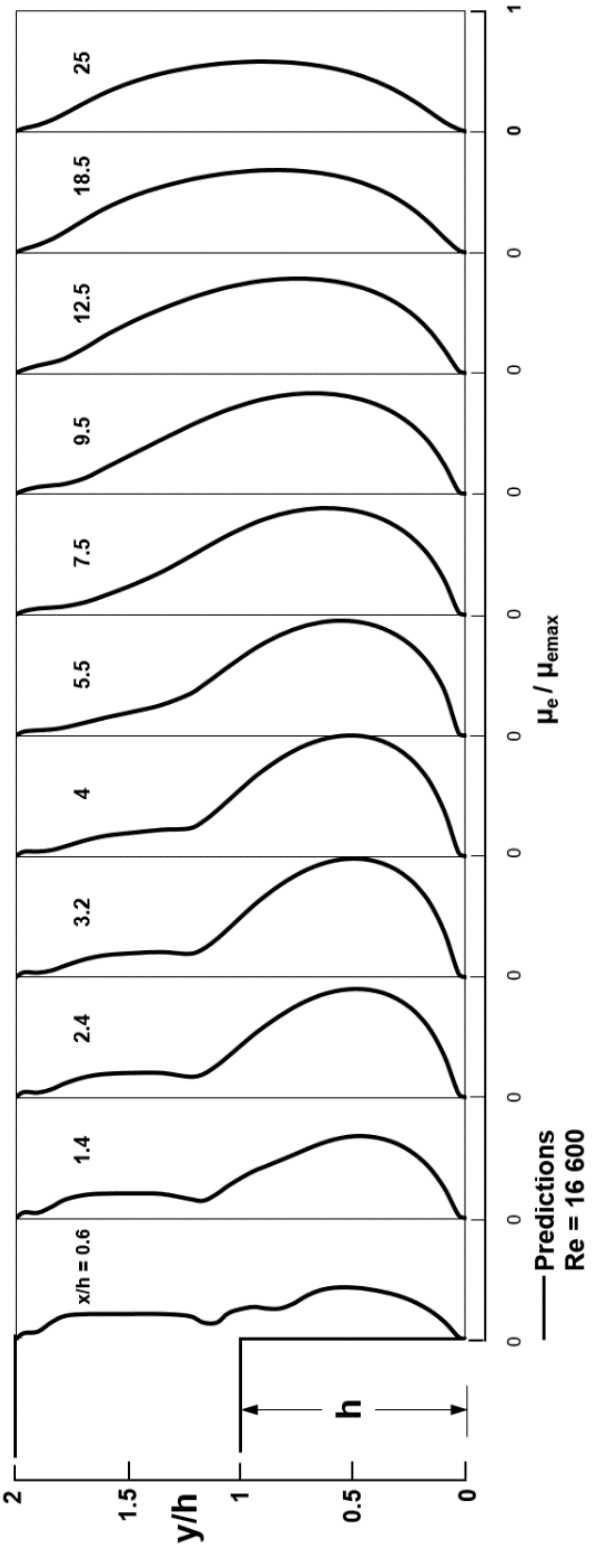


Figure 31: Predicted profiles of normalised effective viscosity along backward-facing step flow channel of [11].

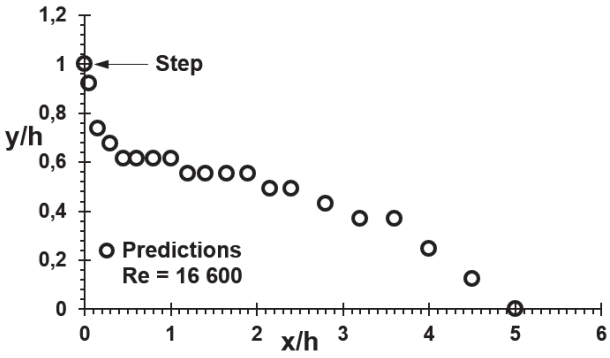


Figure 32: Predicted locus of flow reversal (u=0) for backward-facing step flow of [11].

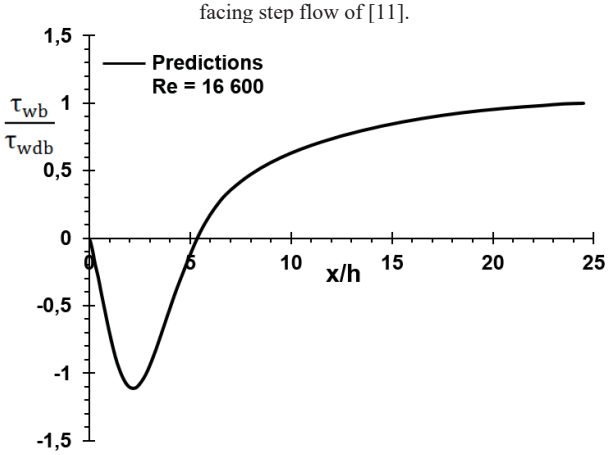


Figure 33: Predicted distribution of dimensionless wall-shear stress along bottom-wall of backward-facing step flow channel of [11].

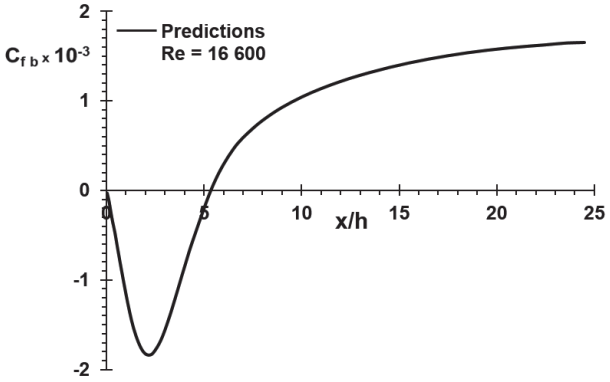


Figure 34: Predicted distribution of friction coefficient along bottom-wall of backward-facing step flow channel of [11].

In Figure 36, the predicted distribution of wall-shear stress along the channel top-wall in dimensionless form of τ_{wt} / τ_{wdt}, as a function of downstream distance x/h, is manifested. Here, the wall-shear stress has been normalised with respect to its value at the outlet plane (τ_{wdt}). As can be seen from the figure, starting from the inlet plane of the channel until at about x/h=0.6 the wall-shear stress increases sharply until it reaches its maximum positive value, and thereafter decreases up to x/h=24.5. Finally, the computed variation of friction coefficient ($C_{ft} = 2\ \tau_{wt} / \rho\ u_o^2$) along the channel top-wall, as a function of downstream distance x/h, is presented in Figure

37. As seen from the figure, the friction coefficient reveals the same trend as the wall-shear stress shown in Figure 36.

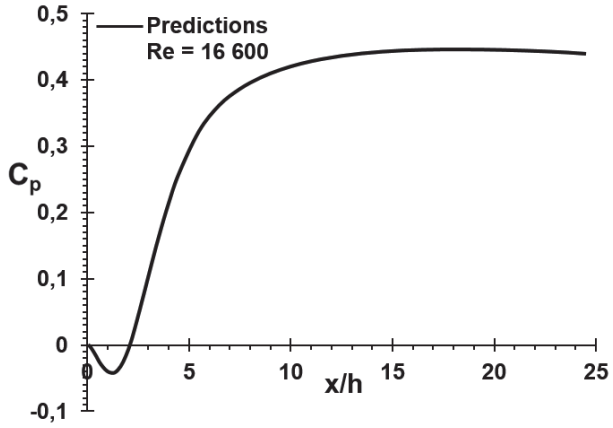


Figure 35: Predicted distribution of wall static-pressure coefficient along bottom-wall of backward-facing step flow channel of [11].

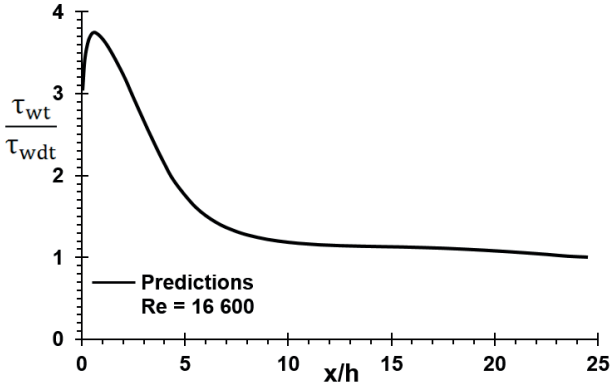


Figure 36: Predicted distribution of dimensionless wall-shear stress along top-wall of backward-facing step flow channel of [11].

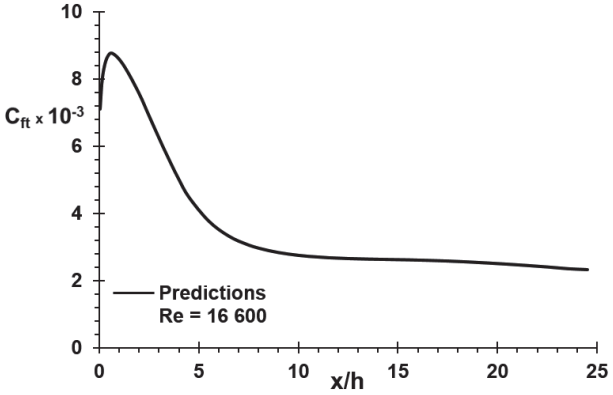


Figure 37: Computed variation of friction coefficient along top-wall of backward-facing step flow channel of [11].

IV. CONCLUDING REMARKS

The main concluding remarks from the numerical study of the present original research work can be summarised as follows. Steady, incompressible, two-dimensional, separating and reattaching complex turbulent recirculating flow over backward-facing steps at three different Reynolds numbers has been computed numerically employing the standard k-ε turbulence model. Using the finite-volume approach, a computer program based on the SIMPLE algorithm of Patankar [46] has been developed. The performance of the standard k-ε turbulence model has been investigated for three different backward-facing step flows. Numerical predictions for local streamwise velocity, turbulence kinetic energy, turbulence kinetic energy dissipation rate, effective viscosity profiles, locus of flow reversal, wall-shear stress, friction coefficient and wall static-pressure coefficient distributions along the backward-facing step flow channels top and bottom walls are presented and, wherever availlable, compared with experimental measurements reported in the literature. The results of numerical study and modelling are generally in good agreement with experimental measurements.

REFERENCES

[1] E W Adams, and J K Eaton, "An LDA study of the backward-facing step flow, including the effects of velocity bias," *ASME J. Fluids Eng* , 110, pp 275-282, 1988

[2] M K Denham, P Briard, and M A Patrick, "A directionally sensitive laser anemometer for velocity measurements in highly turbulent flows," *J. Physics E: Scientific Instruments*, 8, pp 681-683, 1975

[3] W H Stevenson, H D Thompson, and R R Craig, "Laser velocimeter measurements in highly turbulent recirculating flows," *ASME J. Fluids Eng* , 106, pp 173-180, 1984

[4] J R Fessler, and J K Eaton, "Particle response in a planar sudden expansion flow," *Experimental Thermal and Fluid Science*, 15, pp 413-423, 1997

[5] J C Vogel, and J K Eaton, "Combined heat transfer and fluid dynamic measurements downstream of a backward-facing step," *ASME J. Heat Transfer*, 107, pp 922-929, 1985

[6] K B Chun, and H J Sung, "Control of turbulent separated flow over a backward-facing step by local forcing," *Experiments in Fluids*, 21, pp 417-426, 1996

[7] J Kim, S J Kline, and J P Johnston, "Investigation of a reattaching turbulent shear layer: flow over a backward-facing step," *ASME J. Fluids Eng* , 102, pp 302-308, 1980

[8] W A de Groot, "Laser doppler diagnostics of the flow behind a backward-facing step," Ph D Thesis, Georgia Institute of Technology, Atlanta, GA, 1985

[9] R V Westphal, and J P Johnston, "Effect of initial conditions on turbulent reattachment downstream of a backward-facing step," *AIAA Journal*, 22(12), pp 1727-732, 1984

[10] B Ruck, and B Makiola, "Particle dispersion in a single-sided backward-facing step flow," *Int. J. Multiphase Flow*, 14, pp 787-800, 1988

[11] M V Ötügen, "Expansion ratio effects on the separated shear layer and reattachment downstream of a backward-facing step," *Experiments in Fluids*, 10, pp 273-280, 1991

[12] N Kasagi, and A Matsunaga, "3-D particle-tracking velocimetry measurement of turbulence statistics and energy budget in a backward-facing step flow," *Int. J. Heat and Fluid Flow*, 16, pp 477-485, 1995

[13] J T Yang, B B Tsai, and G L Tsai, "Separated-reattaching flow over a backstep with uniform normal mass bleed," *ASME J. Fluids Eng* , 116, pp 29-35, 1994

[14] D M Driver, and H L Seegmiller, "Features of a reattaching turbulent shear layer in divergent channel flow," *AIAA Journal*, 23(2), pp 163-171, 1985

[15] T Karasu, *Numerical prediction of incompressible turbulent swirling flows in circular-sectioned ducts and annuli*, Ph D Thesis, Imperial College of Science, Technology and Medicine, University of London, London, United Kingdom, 1980

[16] T Karasu, P R Choudhury, and M Gerstein, "Prediction of some turbulent flows using upwind and hybrid discretisation schemes and the two-equation turbulence model," *Proc. 4th Miami international symposium on multi-phase transport and particulate phenomena*, Miami Beach, Florida, U S A , 5, pp 105-124, 1988

[17] T Karasu, "Prediction of turbulent flow with heat transfer in an annulus with rotating inner cylinder," *Proc. 8th Miami international conference on alternative energy sources*, Miami Beach, Florida, U S A , vol 2, pp 361-385, 1989

[18] T Karasu, "Numerical prediction of turbulent swirling flows in circular-sectioned annuli," *Proc. 5th Miami international symposium on multi-phase transport and particulate phenomena*, Miami Beach, Florida, U S A , vol 1, pp 149-174, 1990
[19] T Karasu, "Numerical study of turbulent flow in pipes," *Proc. 6th Miami international symposium on heat and mass transfer*, Miami, Florida, U S A , 1990
[20] T Karasu, "Numerical prediction of turbulent recirculating flow through axisymmetric sudden expansions," *10th International Conference on Numerical Methods in Laminar and Turbulent Flow*, 21st –25th July 1997, Swansea, U K , Proceedings Book, Vol 10, pp 357-368, 1997
[21] T Karasu, "Numerical computation of turbulent flow in pipes," *Doğa-Tr. J. of Engineering and Environmental Sciences,* 17, pp 29-38, 1993
[22] T Karasu, "Numerical solution of turbulent flow with heat transfer in an annulus with rotating inner cylinder," *Turkish Journal of Engineering and Environmental Sciences*, Vol 12, No 3, pp 250-272, 1988
[23] T Karasu, "Numerical prediction of turbulent flow in circular pipes," *9th International Conference on Numerical Methods in Laminar and Turbulent Flow*, Atlanta, Georgia, U S A , Proceedings Book, Vol 9, Part 2, pp 1329-1339, 1995
[24] T Karasu, "Kapalı türbülanslı akışların modellemesi ve bilgisayarlı simülasyonu," *20. Ulusal Mekanik Kongresi*, 05-09 Eylül 2017, Uludağ Üniversitesi, Bursa, Türkiye Bildiriler Kitabı, s 1-12, 2017 ISBN 978-975561491-5
[25] T Karasu, "Numerical investigation and modelling of confined turbulent recirculating flows," *International Conference on Engineering Technologies (ICENTE'18)*, 26-28 October 2018, Selçuk University, Konya, Turkey, Proceedings Book, pp 419-433, 2018 E-ISBN: 978-605-68537-3-9
[26] T Karasu, "Computational investigation of turbulent flow in single-sided backward-facing steps," *International Conference on Engineering Technologies (ICENTE'19)*, 25-27 October 2019, Selçuk University, Konya, Turkey, Proceedings Book, pp 321-334, 2019 E-ISBN: 978-605-68537-9-1
[27] T Karasu, "Computational investigation of turbulent flow in pipes," *3rd International Conference on Advanced Technology and Sciences (ICAT'16)*, 01-03 September 2016, Selçuk University, Konya, Turkey, Proceedings Book, pp 939-948, 2016
[28] T Karasu, "Numerical computation of developing turbulent flow between two parallel plates," *21. Ulusal Isı Bilimi ve Tekniği Kongresi (ULIBTK'17)*, 13-16 Eylül 2017, Hitit Üniversitesi, Çorum, Türkiye, Bildiri Kitabı, ss 11-20, 2017
[29] T Karasu, "Numerical study of turbulent flow in circular-sectioned pipes," *8th Atmospheric Sciences Symposium (ATMOS'2017)*, 01-04 November 2017, Istanbul Technical University, Istanbul, Turkey, Proceedings Book, pp 663-674, 2017 ISBN**:** 978-975-561-490-8
[30] T Karasu, "Numerical prediction and modelling of internal turbulent flows with recirculation," *9th Atmospheric Sciences Symposium (ATMOS'2019)*, 23-26 October 2019, Istanbul Technical University, Istanbul, Turkey, E-Proceedings Book, pp 51-64, 2019 e-ISBN: 978-975-561-482-3
[31] T Karasu, "Numerical investigation of turbulent recirculating flow through plane symmetric sudden expansions," *IV. Uluslararası Katılımlı Anadolu Enerji Sempozyumu (AES'2018)*, 18-20 Nisan 2018, Trakya Üniversitesi, Edirne, Türkiye, Bildiri Kitabı, part 3, ss 2044-2056, 2018
[32] T Karasu, "Computer simulation and modelling of turbulent flow over backward-facing steps," *5th International Conference on Advances in Mechanical Engineering (ICAME'2019)*, 17-19 December 2019, Yıldız Technical University, Istanbul, Turkey, Proceedings Book, pp 803-817, 2019 ISBN 978-605-9546-16-4
[33] T Karasu, "Kapalı türbülanslı çevrintili akışların modellemesi ve sayısal araştırılması," *14. Ulusal Tesisat Mühendisliği Kongresi (TESKON 2019)*, 17-20 Nisan 2019, MMO Tepekule Kongre ve Sergi Merkezi, İzmir, Bildiriler Kitabı, ss 1276-1297, 2019 MMO Yayın No: E/MMO/702-1 978-605-01-1270-2
[34] T Karasu, "Numerical simulation of internal turbulent flows," *International Conference on Engineering Technologies (ICENTE'17)*, 07-09 December 2017, Selçuk University, Konya, Turkey, Proceedings Book, pp 679-688, 2017
[35] T Karasu, "Computer simulation and modelling of cofined turbulent flows," *21. Ulusal Isı Bilimi ve Tekniği Kongresi (ULIBTK'17)*, 13-16 Eylül 2017, Hitit Üniversitesi, Çorum, Türkiye, Bildiri Kitabı, ss 1-10, 2017
[36] T Karasu, "Numerical analysis and modelling of internal turbulent flows with recirculation," *22nd Congress on Thermal Science and Technology (ULIBTK 2019)*, 11-14 September 2019, Kocaeli University, Kocaeli, Turkey Edited Book of Proceedings (Seçilmiş Bildiriler Kitabı), pp 172-187, 2019
[37] T Karasu, "Borularda türbülanslı akışların sayısal simülasyonu," *VIII. Ulusal Hidrolik Pnömatik Kongresi,* 22-25 Kasım 2017, İzmir Bildiriler Kitabı, ss 69-83, 2017 MMO Yayın No: E/MMO/678 ISBN 978-605-01-1088-3
[38] T Karasu, "Paralel iki plaka arasında gelişen türbülanslı akışın sayısal araştırılması," *20. Ulusal Mekanik Kongresi*, 05-09 Eylül 2017, Uludağ Üniversitesi, Bursa, Türkiye Bildiriler Kitabı, ss 13-24, 2017 ISBN 978-975561491-5
[39] T Karasu, "Dahili türbülanslı çevrintili akışların modellemesi ve sayısal hesaplanması," *21. Ulusal Mekanik Kongresi*, 02-06 Eylül 2019, Niğde Ömer Halisdemir Üniversitesi, Niğde, Türkiye Bildiriler Kitabı, ss 1-23, 2019 ISBN 978-975561506-6
[40] T Karasu, "Computer simulation of turbulent recirculating flow through circular-sectioned sudden expansion pipes," *7th International Conference on Advanced Technologies (ICAT'18)*, April 28-May 1, 2018, Antalya, Turkey, Proceedings Book, pp 752-763, 2018 E-ISBN: 978-605-68537-1-5
[41] T Karasu, "Numerical investigation of turbulent recirculating flow in double-sided planar sudden expansions in channels," *7th International Conference on Advanced Technologies (ICAT'18)*, April 28-May 1, 2018, Antalya, Turkey, Proceedings Book, pp 764-773, 2018 E-ISBN: 978-605-68537-1-5
[42] T Karasu, "Computer analysis and modelling of confined turbulent recirculating flows," *4th International Conference on Advances in Mechanical Engineering (ICAME'2018)*, 19-21 December 2018, Yıldız Technical University, Istanbul, Turkey, Proceedings Book, pp 363-382, 2018 ISBN 978-605-9546-13-3
[43] T Karasu, O Kural, N Eğrican, O Borat and O Özcan, "Numerical simulation of turbulent flow in axisymmetric sudden expansions," *Fourth Annual Mechanical Engineering Conference of ISME and Second International Mechanical Engineering Conference*, 14-17 May 1996, School of Engineering, Shiraz University, Shiraz, Iran, Proceedings Book, vol 2, pp 261-268, 1996
[44] T Karasu, "Computational study of turbulent flow in circular pipes," *International Congress on Computational Methods in Engineering*, 2-6 May 1993, School of Engineering, Shiraz University, Shiraz, Iran, Proceedings Book, vol 2, pp 3-10, 1993
[45] B E Launder, and D B Spalding, "The numerical computation of turbulent flows," *Comp. Meth. Appl. Mech. Engng*, 3, pp 269-289, 1974
[46] S V Patankar, *Numerical heat transfer and fluid flow*, Hemisphere, McGraw-Hill, Washington, D C , Chapters 5 and 6, 79-138, 1980
[47] S V Patankar, and D B Spalding, "A calculation procedure for heat, mass and momentum transfer in three-dimensional parabolic flows," *Int. J. Heat Mass Transfer*, 15, pp 1787-1806, 1972
[48] D B Spalding, "A general-purpose computer program for multi-dimensional one-and-two phase flow," *Math. Comput. Simulation*, XXIII, pp 267-276, 1981

Effect of Hydrothermal Pretreatment on Fuel Properties of Coal-Biomass Blend

S. BAŞAKÇILARDAN KABAKCI[1] and M. H. TANIŞ[2]

[1] Yalova University, Engineering Faculty, Energy Systems Engineering Department, Yalova/Turkey, sibel.kabakci@yalova.edu.tr

[2] Yalova University, Institute of Science, Energy Systems Engineering Department, Yalova/Turkey, tansmedyahatun@gmail.com

Abstract **- Hydrothermal pretreatment is an effective method used to improve the fuel properties of both lignocellulosic biomass sources and low-quality coals. In the present study, firstly, hydrothermal pretreatment was applied to wood sawdust and Soma lignite individually, and the optimum conditions (reaction temperature and reaction period) were determined. Then, hydrothermal pretreatment was applied to blend consisting of 50% Soma lignite and wood sawdust by weight under the determined optimum conditions (90 min, 230 °C). After hydrothermal pretreatment, the heating value increased both in the raw samples (Soma lignite and wood sawdust) and in the blend. Based on results, the highest increase in heating value was 22.7% (90 min, 230 °C) for wood sawdust and 10.3% (90 min, 220 °C) for lignite. While only dehydration was observed during hydrothermal treatment of raw samples, decarboxylation was also seen along with dehydration in the Soma lignite-wood sawdust blend. After hydrothermal pretreatment, while ash, volatile matter, hydrogen and oxygen content decreased, carbon content increased in all samples. The hydrothermally pretreated lignite-wood sawdust mixture (50% by weight) had 13.20% higher heating value and 19.65% less ash compared to the mixture of individually hydrothermally treated wood sawdust and lignite (50% by weight). It was found that the ignition temperature and the burn-out temperature of the hydrothermally treated lignite-wood sawdust blend were between the respective temperatures of the raw materials forming the blend. The ignition temperature of the hydrothermally pretreated lignite-wood sawdust blend increased by 7 °C compared to the raw blend and the burn-out temperature decreased by 10 °C compared to the raw blend.**

Keywords **- Hydrothermal treatment, Low-quality lignite, Coal-biomass blends, Combustion characteristics**

I. INTRODUCTION

According to the report (Coal 2018-Analysis and Forecasts to 2023) released by the International Energy Agency in December 2018, coal met 27% of all energy consumed in the world and 38% of electricity production [1]. This is mainly because of that; coal plays a very important role in energy production as well as in sectors such as iron and steel. Unlike bituminous coals, low-quality coals (lignite and sub-bituminous coals) are more widely distributed around the world [2], and their reserves are about 500 billion tons. Reducing coal-based electricity production in developed countries as much as possible is an important policy to combat climate change. However, even though its contribution to climate change is known, coal is still preferred in developing countries where the energy security is at risk. The reason is that coal is an abundant, local and economic resource in those countries.

It is an effective method to use biomass together with coal in order to reduce the environmental effects that they create as a result of combusting fossil fuels [3, 4]. It is known that when biomass is used instead of coal at certain rates, fossil-based carbon dioxide emissions reduce approximately 1.05 tons of fossil CO_2 per tonne of woody biomass burned [5]. The use of biomass together with coal also causes a decrease in SO_x emissions and reduces the fuel cost [6]. Another advantage of mixing biomass with coal is that it increases the reactivity of fuel due to the oxygen and hydrogen content that biomass has. However, high moisture and alkali metal content and low burn-out temperature often limit the use and percentage of biomass in biomass-coal mixtures [3].

Hydrothermal treatment is an important pretreatment that improves the fuel properties of both low quality coals and biomass. Hydrothermal treatment is generally carried out under inert atmosphere, at low temperatures (180-280 °C) and under pressure to keep the water in liquid form [7]. With hydrothermal pretreatment, the moisture and volatile content of the biomass are reduced and the energy density is increased. Moreover, inorganics in the structure of the biomass dissolve in the liquid phase during the hydrothermal treatment. Therefore, it is possible to state that hydrothermally pretreated biomass will have less tendency to form slag [8]. Hydrothermal treatment is also applied to low quality coals [2]. Low-quality coals are mainly recognized by their low calorific value and carbon content. Besides, they have also high moisture (>25%) and oxygen contents (> 14%) [9]. Hydrothermal pretreatment, which is applied to low quality coals, is mostly done for dewatering [10, 11]. In the study of Mingshun et al. [12], it has been shown that moisture, oxygen and sulfur content in hydrothermally treated sub-bituminuous coals decreased significantly and the heating value increased. In the same study, it was stated that the alkali metal content in the coal also decreased, which resulted in lower tendency for slag formation.

In the literature, it is seen that hydrothermal treatment has been applied to different biomass sources and different quality coals separately. There are also studies where hydrothermally treated biomass is mixed with low-quality coal or where separately hydrothermally treated coal and biomass are used by mixing. In a similar study by Parshetti et al. [13], hydrothermally treated sewage sludge was mixed with low-quality Indonesian coal and hydrothermally treated low-quality Indonesian coal separately. Their study showed that co-combustion of hydrotreated sewage sludge with either low quality Indonesian coal or hydrotreated low quality Indonesian coal could improve the quality of emissions. Since the hydrothermal treatment causes separate changes in the structure of biomass and coal, a co-hydrothermal treatment to its mixtures can result in a better fuel compared to the raw mixture. Saba et al. [14] performed a similar study (mischantus and coal were co-treated hydrothermally) and showed the synergistic effect of biomass on bituminous coal during hydrothermal treatment. The synergistic effect was also seen in the study of Nonaka et al. [15], in which polymerization was stimulated to form larger molecules during the co-hydrothermal process of low-quality coal and biomass.

In this study, it was investigated how hydrothermal pretreatment affected fuel properties and combustion characteristics of wood sawdust, lignite and their blend (50% by weight). For this purpose, firstly, the optimum hydrothermal pretreatment conditions (temperature and reaction period) that increase the heating value were determined, and then hydrothermal pretreatment was applied to wood sawdust, Soma lignite and their blend (50% by weight). Finally, how the fuel properties and the combustion characteristics changed after hydrothermal treatment have been discussed.

II. Materials and Method

A. Materials

Poplar wood sawdust originated from Ortaca (Muğla) was used as the source of lignocellulosic biomass. Soma lignite was obtained from the Soma Basin (Manisa). All raw materials were grounded and passed through a 250 μm (60 mesh) sieve prior to hydrothermal treatment.

B. Method

In this study, firstly, optimum conditions of hydrothermal pretreatment were determined. For this, two variables (temperature and reaction period) were examined at three levels. The temperature levels were chosen as 200, 220 and 230 °C. Another variable, reaction period, was fixed as 30, 60 and 90 min separately. In the current study, the raw material/water ratio was kept constant at the ratio of 1:4 (g/ml), taking into account the safe operating limit of the reactor. The solid-liquid mixture was placed in a stainless-steel autoclave. After hydrothermal treatment, the autoclave was cooled, and the mixture was filtered. The solid samples were dried at 105 °C for 24 h. The effectiveness of the pretreatment was evaluated by the increase in the heating value of the hydrothermally treated samples.

C. Analysis

For all samples, moisture content was determined by using American Society for Testing and Materials (ASTM) E871-82 method [16] based on TG analysis; volatile matter was determined by using ASTM E872-82 method [17]; ash content was determined by using ASTM E1755-01 method [18]. The fixed carbon content was calculated from the difference. An elemental analyzer (LECO/TRUSPEC MICRO) was used for determining the C, H, N and S samples. IKA C5003 model bomb calorimeter was used for determining the heating value. The functional groups were identified by Perkin Elmer Spectrum 100 Fourier Transform Infrared Spectroscopy (FTIR). The spectra were recorded with a resolution of 4 cm^{-1} in the range from 400 to 4000 cm^{-1}.

Combustion characteristics of the samples were obtained by using a thermogravimetric analyser (SEIKO, TG/DTA 6300) under a constant air flow (200 mL/min). Each sample was heated to 105 °C (heating rate: 10 °C/min), and kept at this temperature for 10 min. Then, it was heated to 900 °C (heating rate: 15 °C/min) and kept at this temperature for 7 min. Ignition temperature, burn-out temperature and the temperature of maximum weight loss rate were determined by using the thermal gravimetry data and the corresponding derivative weight loss data (DTG). For finding the ignition temperature and burn-out temperature, intersection method was used [19].

III. Results and Discussion

A. Hydrothermal pretreatment conditions

Figure 1 shows the percentage increase in the heating value of the hydrothermally treated samples.

According to the results of the experiments with wood sawdust, higher temperature and reaction periods resulted with increased the heating value. The effect of temperature can be explained by the fact that temperature directly affect water properties in the sub-critical conditions [20]. Apart from affecting solubility, temperature also affects the decomposition of hemicellulose during the hydrothermal treatment. Another important factor, the reaction period, was shown to form polyaromatic structures of monomers dissolved in the liquid phase [21]. In this experimental study, treatment period was observed as a more effective parameter than the temperature at which the treatment was carried out. The percentage increase in heating value of all poplar sawdust hydrochars was between 1.3 % (30 min, 230 °C) and 22.7 % (90 min, 230 °C). The effect of conditions on hydrothermal treatment was also observed by Nakason et al. [22], who treated coconut husk and rice husk at low temperatures (140 °C -200 °C) and wide reaction periods (1 h-4 h). It was observed that the reaction period was a more effective parameter than the temperature in hydrothermal carbonization of both lignocellulosics.

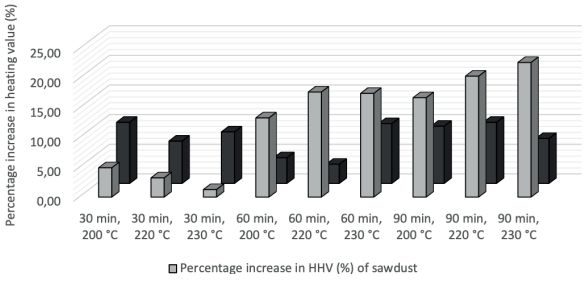


Figure 1. The percentage increase in the heating value under different pretreatment conditions.

Unlike poplar sawdust, it was observed that the effect of temperature and treatment period on heating value of hydrothermally treated Soma lignite was very low. The increase in the heating value of the hydrothermally treated Soma lignite was between 3.2 % (60 min, 220 °C) and 10.3 % (90 min, 220 °C). Similar observation was done by Zhang et al. [23], who showed that there was a slight increase in the heating value of coal as the hydrothermal treatment temperature was increased. In another study done by Feng et al. [24], hydrothermal treatment at 300 °C was observed as effective compared to higher temperatures. Higher temperatures may promote decarboxylation reactions to occur during hydrothermal treatment. In the study done by Zheng et al. [25], it was shown that reaction periods over 60 min did not affect the elemental composition of coal. In the same study, it was observed that the efficiency decreased in hydrothermal processes above 350 °C. It is quite clear that coal type is also an important factor to decide on the conditions of the pretreatment.

Based on the experimental results, optimum period for the co-hydrothermal treatment was selected as 90 min, and the optimum temperature of the co-treatment was determined as 230 °C.

B. Effect of hydrothermal pretreatment on fuel properties

After hydrothermal pretreatment, the changes in the approximate analysis and elemental analysis of poplar wood sawdust, Soma lignite and the blend (50 % by weight) are shown in Table 1. In terms of ash and carbon content, the difference between poplar sawdust and Soma lignite was remarkable. As observed in proximate and elemental analysis, poplar sawdust had high volatile matter and high oxygen content, like most lignocellulosics. Soma lignite, like low-rank coals, had high ash content and very low fixed carbon content. Although the fuel properties of poplar sawdust and Soma lignite were quite different from each other, the effect of hydrothermal treatment on these materials was similar. Hydrothermal treatment resulted with higher fixed carbon content and lower volatile matter and ash content in both poplar sawdust and Soma lignite. Since the hydrolysis of hemicellulose occurs under hydrothermal treatment conditions, the rate of decrease in volatile matter content of poplar sawdust was higher than that of Soma lignite. The percentage of decrease in volatile matter was 8.31% for poplar sawdust and 2.94% for Soma lignite. The volatile matter content of the co-hydrothermally treated blend was 32.7% less than the volatile content of the raw mixture.

Van Krevelen diagram showing the relation between O/C ratio and H/C ratio of raw and hydrothermally treated samples is given in Figure 2. In the Van Krevelen diagram, the biomass is usually in the upper right corner of the diagram; lignite, bituminous coal and anthracite move to the lower left corner of the diagram. The intensity of the hydrothermal treatment depends on how the atomic proportions (O/C and H/C) change. Hydrothermal treatment includes dehydration, decarboxylation and demethanation reactions [26]. According to the experimental results, the H/C and O/C atomic ratios of all solid products obtained after hydrothermal treatment were decreased. As seen in the Van Krevelen diagram, the major reaction observed in hydrothermal treatment of polar sawdust and Soma lignite was dehydration [27, 28]. In the case of co-hydrothermally treated blend, decarboxylation was also seen along with dehydration. This was also noticeable by the fact that the increase in carbon percentage after hydrothermal treatment was lower in the blend (13.19%) than poplar sawdust (22.11%) and Soma lignite (13.87%).

Table 1: Proximate and elemental analysis of poplar sawdust, Soma lignite, poplar sawdust-Soma lignite blend (50%, wt. %) and their hydrotreated samples

	Proximate analysis (wt. %, db)			Elemental analysis (wt.%, db)				
	Ash	VM	FC	C	H	N	S	O
PS	5.45	72.44	22.11	46.26	6.09	0.23	-	47.42
HPS	4.24	66.42	29.34	56.49	5.76	0.35	-	37.4
SL	66.88	19.76	13.36	17.01	1.61	0.33	0.28	80.77
HSL	64.3	19.18	16.52	19.37	1.59	0.47	0.21	78.36

PS: Poplar sawdust, HPS: Hydrotreated poplar sawdust
SL: Soma lignite, HSL: Hydrotreated Soma lignite

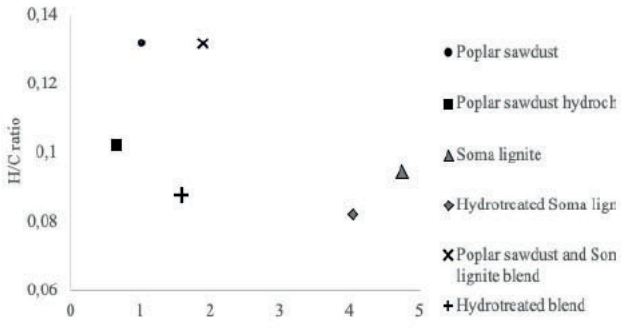


Figure 2: Van Krevelen diagram of raw and hydrotreated samples.

Figure 3-5 show the FTIR spectra of the raw and hydrotreated samples. O-H stretching vibration of hydroxyl group was observed at 3500-3300 cm^{-1} in both poplar sawdust and its hydrochar (Figure 2) [22]. Due to dehydration, O-H stretching vibration was less in hydrochar. Similarly, aliphatic C-H vibration at 2900cm^{-1} was also observable in both poplar sawdust and its hydrochar [29]. Carboxylic acid groups of hemicellulose at 1732 cm^{-1} and aromatic C-H group of hemicellulose at 895cm^{-1} were detectable in raw poplar sawdust while they were not observed in the structure of its hydrochar. Degradation of hemicellulose at hydrothermal conditions was the main reason of this disappearance. C=O stretching vibration at 1694cm^{-1}, aromatic skeletal vibration derived from lignin at 1602cm^{-1} [30], guaiacyl unit of lignin at 1266cm^{-1} [31] were intensified in the hydrochar sample.

As seen in Figure 3, hydrothermal treatment did not change the structure of the Soma lignite at all. According to Nonaka et al. [15], hydrothermal treatment did not cause major changes in the structure of coal. In the current study, the stretching vibration of the hydroxyl groups in clay minerals at 3695 cm^{-1} and 3621 cm^{-1} was observed in both samples. The wide band seen due to OH vibration at 3399 cm^{-1} was observed only in the raw Soma lignite. The stretching vibration of carboxyl group at 1622 cm^{-1}, stretching vibration of C=C at 1428 cm^{-1} and Si-O-Si vibration at 1000 cm^{-1} were observable in both raw and hydrotreated lignite sample [9].

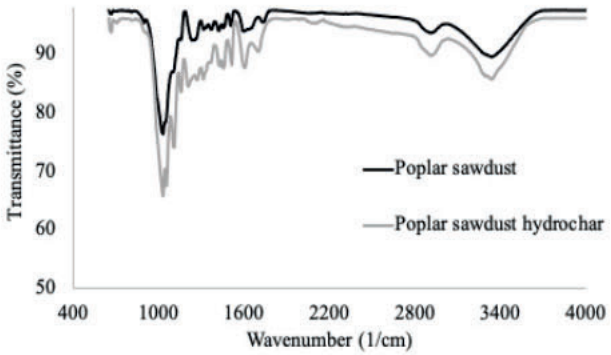


Figure 3: FTIR spectra of poplar sawdust and its hydrochar.

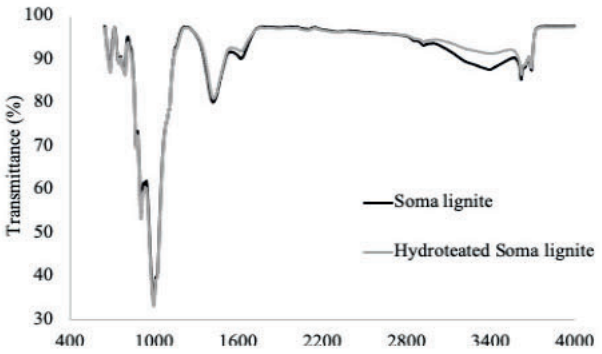


Figure 4: FTIR spectra of Soma lignite and its hydrothermally treated sample.

The FTIR spectra of the blend (see in Figure 4) showed the structural groups representing the both raw materials. FTIR spectra of the raw blend indicated the Si-O-Si vibration (1000 cm^{-1}), stretching vibration of aromatic C=C groups (1428 cm^{-1}) and -OH groups in clay minerals (3695 cm^{-1}), which were all originated from raw Soma lignite. Similarly, vibrations belonging to the specific structural groups of poplar sawdust were seen in the FTIR spectra of the raw blend. Aliphatic C-H vibration (2900 cm^{-1}) and aromatic skeleton vibration of lignin (1602 cm^{-1}) were quite remarkable. The FTIR spectra of the hydrotreated blend were a bit different from those of the raw blend. C-O-C stretching in cellulosic structures (1030 cm^{-1}) was quite evident in the FTIR spectra of raw poplar sawdust, raw blend and hydrothermally treated blend. While the aromatic skeletal vibration of lignin (1513 cm^{-1}) did not appear very clearly in the FTIR spectra of the raw poplar sawdust, it was observed in the FTIR spectra of both the poplar sawdust hydrochar and the hydrothermally treated blend. Also, the C=O stretching vibration at 1700 cm^{-1} was observable in the spectra of both poplar sawdust hydrochar and hydrotreated blend.

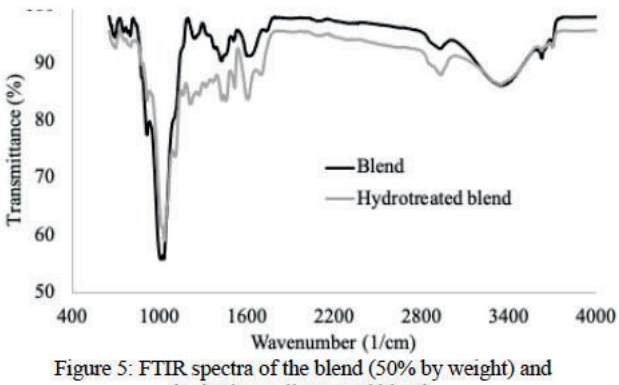


Figure 5: FTIR spectra of the blend (50% by weight) and hydrothermally treated blend.

C. Combustion characteristics of raw and hydrothermally treated samples

Combustion profile of wood sawdust (see in Figure 5) involves the combustion steps of a typical biomass: moisture removal, devolatilization-combustion of volatiles and oxidation of char. According to the DTG curve, in the combustion process, moisture was removed first (<100 °C), then the volatiles formed were burned (260-359 °C) and char was oxidized between 419-477°C. The most important difference in poplar hydrochar combustion is the delay of the initial temperature of devolatilization. This was because hemicellulose was degraded under hydrothermal conditions. The degradation of hemicellulose also caused the volatile material to decrease. Due to this decrease, the formation and burning of volatiles were delayed to 298 °C and 355 °C, and oxidation of the char was between 389 °C and 497 °C. The temperature at which the maximum mass loss rate was recorded was unchanged and was determined as 333 °C.

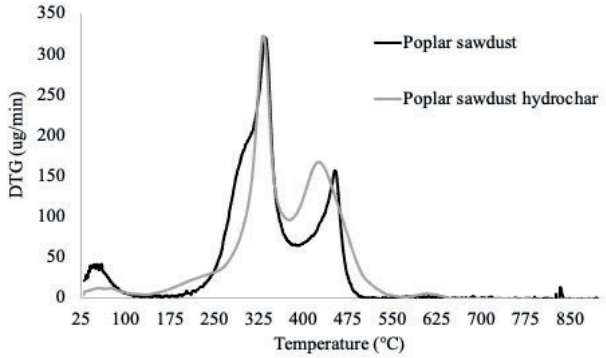


Figure 6: Combustion profile of poplar sawdust and its hydrochar.

Compared to poplar sawdust, Soma lignite needed higher temperatures for devolatilization, combustion of volatiles and oxidation of char (see in Figure 7). It was observed that the combustion of Soma lignite occurred in a wide temperature range (350-719 °C). The stage including devolatilization, combustion of the volatiles and oxidation of the char was observed at temperatures between 350-500 °C. This stage caused a mass loss of about 22 %. Although the combustion profile of the hydrothermally treated Soma lignite was similar to that of the raw Soma lignite, it was observed that the temperatures corresponding to the burning stages were a bit different. Combustion process of hydrothermally treated Soma lignite has also been observed to occur in a wide temperature range (326-729 °C). While the temperature at which the maximum mass loss was observed during the combustion process of the raw Soma lignite was 407 °C, this temperature decreased to 394 °C in the hydrothermally treated Soma lignite. The decomposition (at temperatures between 625-775 °C) of minerals such as calcite and kaolinite in the structure of coal caused a significant loss of mass (8.7 % mass loss of raw Soma lignite, 10.3 % mass loss of hydrotreated Soma lignite).

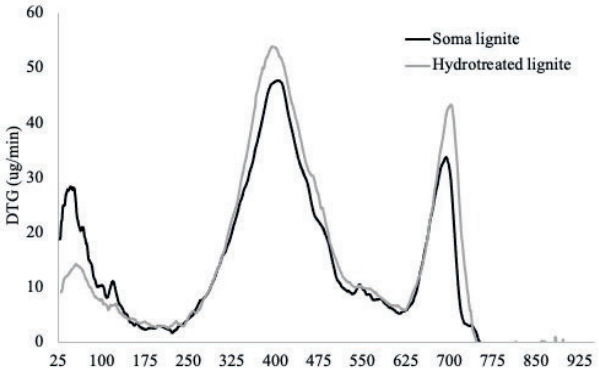


Figure 7. Combustion profile of lignite and hydrothermally treated lignite.

The combustion profile of the raw blend (see in Figure 8) was like the combination of the profile of both poplar sawdust and Soma lignite. Devolatilization, combustion of volatiles and oxidation of char (237.5-558 °C) caused about 59.5 % mass loss. In the combustion profile of the hydrothermally treated poplar sawdust-Soma lignite blend, it was observed that the combustion occurred at a wide temperature range between 250-500 °C. The decomposition of minerals in the structure of Soma lignite was also observed in the raw blend (616-700 °C) and hydrothermally treated blend (630-690 °C).

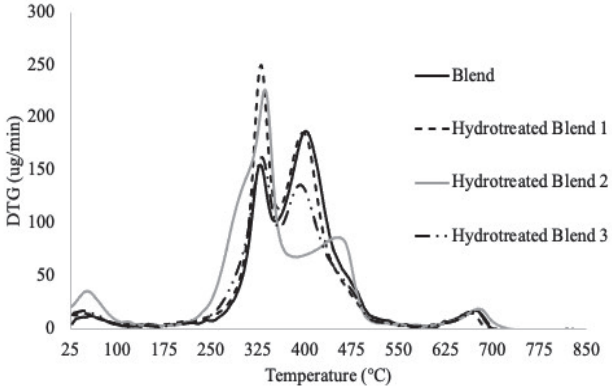


Figure 8. Combustion profile of the blend (50% by weight) and hydrothermally treated blend

Combustion characteristics (ignition temperature, burnout temperature and the temperature of maximum weight loss), which can be determined with thermogravimetric analyzer, can be used to compare the combustibility and reactivity of fuels [32]. In this study, intersection method was used for determining the ignition temperature and burn-out temperature.

Table 2 shows the ignition temperature (T_i), burn-out temperature (T_b) and the temperature of maximum weight loss rate (T_{max}). The ignition temperature depends on how early the devolatilization begins and how quickly the heat is released by the combustion of the volatiles. As seen in Table 2, the

ignition temperature of raw poplar sawdust and Soma lignite were quite different from each other. Typically, biomass has a lower ignition temperature than coal. Because, the volatile matter content of biomass is higher compared to coal. This pulls the ignition temperature to lower temperatures. Since the amount of volatile material was considerably reduced due to the decomposition of hemicellulose during hydrothermal treatment, the ignition temperature of poplar sawdust hydrochar was shifted to higher temperature. Although there was not a dramatic change in the structure of Soma lignite, the ignition temperature was a bit lower than the raw lignite since there was dehydration. In the case of blend, the ignition temperature of the co-hydrothermally treated blend was 7 °C higher than the raw sample. The increase might indicate that combustion was difficult, but low moisture and high volatile matter content would facilitate combustion [11]. Moreover, high ignition temperature is a required property due to the reason that it reduces the risk of spontaneous ignition. This makes to store and transport the upgraded blend safer.

When it comes to burn-out temperature, the difference between biomass and coal was quite evident. Poplar sawdust hydrochar and hydrothermally treated Soma lignite had higher combustion temperature compared to their raw samples. It was also observed by Ullah et al. [11] that the burn-out temperature shifted to higher temperature after hydrothermal treatment. The co-hydrothermal treatment reduced the burn-out temperature of the blend. In other words, the complete combustion of the hydrothermally treated mixture was done in a shorter time.

Table 2. Combustion characteristics of samples

	Poplar sawdust		Soma Lignite		Blend	
	Raw	HTT	Raw	HTT	Raw	HTT
T_i (°C)	277	298	333	326 5	300	307
T_b (°C)	476 7	497	718 7	728 6	700	690
T_{max} (°C)	335 4	332 8	407	393 5	401 7	329 6

HTT: Hydrothermally treated

IV. Conclusion

The use of low-quality lignite in energy production is highly disadvantageous due to low energy density and environmental problems they cause. If low quality coal is desired to be used anyway, mixing with biomass may be a good option. On the other hand, it is possible to further improve the fuel properties of biomass-coal blends. Hydrothermal treatment can be considered as an alternative method for both fuels. The current study also showed hydrothermal treatment might be a good option for upgrading fuel properties of biomass-coal blends. The co-hydrothermal treated mixture had 13.20% more heating value and 19.65% less ash than the mixture of individually hydrothermally treated poplar sawdust and lignite. It was found that the ignition temperature and the burn-out temperature of the co-hydrothermal treated blend were between the respective temperatures of the raw materials forming the blend. The co-hydrothermal treatment increased the ignition temperature of the fuel by 7 °C and decreased the burn-out temperature by 10 °C.

Acknowledgment

This research was financially supported by Yalova University Scientific Research Unit [Project No: 2018/YL/0015].

V. References

[1] IEA (2018), "Coal 2018", IEA, Paris https://www.iea.org/reports/coal-2018

[2] Yu, J., Jiang, C., Guan, Q., Gu, J., Ning, P., Miao, R., & Chen, Q. A. "Conversion of low-grade coals in sub-and supercritical water: A review", *Fuel*, 275-284, 2018.

[3] Loeffler, D. and Anderson, N. "Emissions tradeoffs associated with cofiring forest biomass with coal: A case study in Colorado, USA.", *Applied Energy*, 113, 67-77, 2014.

[4] Gil, M.V. and Rubiera, F. "Coal and biomass cofiring: fundamentals and future trends". (Chapter 5) In: Suarez-Ruiz, I., Diez, M.A., Rubiera, F., ed. *New trends in coal conversion. Combustion, Gasification, Emissions, and Coking*. Woodhead Publishing 117-140, 2019.

[5] Tillman, D.A., Duong, D.N.B., Harding, S. "Blending coal with biomass: Cofiring biomass with coal" (Chapter 4). In: Tillman, D.A., Duong, D.N.B., Harding, S., ed. *Solid fuel blending*. Butterworth-Heinemann; 125-200, 2012.

[6] Adams, P., Bridgwater, T., Lea-Langton, A., Ross, A., Watson, I. "Biomass conversion technologies". (Chapter 8) In: Thornley, P. and Adams, P., ed. *Greenhouse gas balances of bioenergy systems.* Academic Press; 107-139, 2018.

[7] Reza, M., Uddin, M., Lynam, J., Hoekman, S., & Coronella, C. "Hydrothermal carbonization of loblolly pine: reaction chemistry and water balance", *Biomass Conversion and. Biorefinery*, 4, 311-321, 2014.

[8] Gao, L., Volpe, M., Lucian, M., Fiori, L., Goldfarb, J.L. "Does hydrothermal carbonization as a biomass pretreatment reduce fuel segregation of coal-biomass blends during oxidation?", *Energy Conversion and Management*, 181, 93-104, 2019.

[9] Yang, M., Xie, Q., Wang, X., Dong, H., Zhang, H., Li, C. "Lowering ash slagging and fouling tendency of high-alkali coal by hydrothermal pretreatment", *International Journal of Mining Science and Technology*, 29, 521-525, 2019.

[10] Liu, X., Wu, X., Wang, J. "Substantial upgrading of a high-ash lignite by hydrothermal treatmen followed by $Ca(OH)_2$ digesiton/acid leaching", *Fuel*, 222, 260-277, 2018.

[11] Ullah, H., Liu, G., Yousaf, B., Ali, M. U., Abbas, Q., Zhou, C., Rashid, A. "Hydrothermal dewatering of low-rank coals: Influence on the properties and combustion characteristics of the solid products", *Energy*, 158, 1192-1203, 2018.

[12] Mingshun, Y., Qiang, X., Xin, W., He, D., Hao, Z., Chunqi, L. "Lowering ash slagging and fouling tendency of high-alkali coal by hydrothermal pretreatment" *International Journal of Mining Science and Technology*, 29 (3), 521-525, 2019.

[13] Parshetti, G.K., Liu, Z., Jain, A., Srinivasan, M.P., Balasubramanian, R. "Hydrothermal carbonization of sewage sludge for energy production with coal", *Fuel*, 111, 201-210, 2013.

[14] Saba, A., Saha, P., Reza, M.T. "Co-hydrothermal carbonization of coal-biomass blend: influence of temperature on solid fuel properties", *Fuel Processing Technology*, 167, 711-720, 2017.

[15] Nonaka, M., Hirajama, T., Sasaki, K. "Upgrading of low rank coal and woody biomass mixture by hydrothermal treatment", *Fuel*, 90, 2578-2584, 2011.

[16] American Society for Testing and Materials (ASTM). (2013). *E871-82 Standard Test Method for Moisture Analysis of Particulate Wood Fuels.*

[17] American Society for Testing and Materials (ASTM). (2013). *e872-82 Standard Test Method for Volatile Matter in the Analysis of Particulate Wood Fuels.*

[18] American Society for Testing and Materials (ASTM). (2015). *E1755-01 Standard Test Method for Ash in Biomass.*

[19] Başakçılardan Kabakcı, S. "Pyrolysis and combustion characteristics and kinetics of wood sawdust and wood sawdust hydrochar", *Environmental Progress & Sustainable Energy*, e13315, 1-9, 2019.

[20] Wang, T., Zhai, Y., Zhu, Y., Li, C., Zeng, G. "A review of hydrothermal carbonization of biomass waste for hydrochar formation: Process conditions, fundamentals, and physicochemical properties", *Renewable and Sustainable Energy Reviews*, 90, 223-247, 2018.

[21] Kang, S., Li, X., Fan, J., Chang, J. "Characterization of hydrochars produced by hydrothermal carbonization of lignin, cellulose, D-xylose, and wood meal", *Ind Eng Chem Res*, 51, 9023-9031, 2012.

[22] Nakason, K., Panyapinyopol, B., Kanokkantapong, V., Viriya-empikul, N., Kraithong, W., Pavasant, P. "Hydrothermal carbonization of unwanted biomass materials: Effect of process temperature and retention time on hydrochar and liquid fraction", *Journal of the Energy Institute*, 91, 786-796, 2018.

[23] Zhang, X.-P., Zhang, C., Tan, P., Li, X., Fang, Q.-Y. C. "Effects of hydrothermal upgrading on the physicochemical structure and gasification characteristics of Zhundong coal", *Fuel Processing Technology*, 172, 200-208, 2018.

[24] Feng, X., Zhang, C., Tan, P, Zhang, X., Fang, Q., Chen, G. "Experimental study of the physicochemical structure and moisture readsorption characteristics of Zhaotong lignite after hydrothermal and thermal upgrading", *Fuel*, 185, 112-121, 2016.

[25] Zheng, Q., Morimoto, M., Fouquet, T., Sato, H., & Takanohashi, T. "Effect of hydrothermal conditions on production of coal organic microspheres", *Fuel*, 234, 1301-1312, 2018.

[26] Wikberg, H, Grönqvist, S., Niemi, P., Mikkelson, A., Siika-aho, M., Kanerva, H., Kasper, A., Tamminen, T. "Hydrothermal treatment followed by enzymatic hydrolysis and hydrothermal carbonization as means to valorize agro- and forest-based biomass residues", *Bioresource Technology*, 235, 70-78, 2017.

[27] Benstoem, F., Becker, G., Firk, J., Kaless, M., Wuest, D., Pinnekamp, J., Kruse, A. "Elimination of micropollutants by activated carbon produced from fibers taken from wastewater screenings using hydrothermal carbonization", *Journal of Environmental Management*, 211, 278-286, 2018.

[28] Ge, L., Feng, H., Xu, C., Zhang, Y., Wang, Z. "Effect of hydrothermal dewatering on the pyrolysis characteristics of Chinese low-rank coals", *Applied Thermal Engineering*, 141, 70-78, 2018.

[29] Zhu, Z. B., Si, B.C., Lu, J.W., Watson, J., Zhang, Y.H., Liu, Z.D. "Elemental migration and characterization of products during hydrothermal liquefaction of cornstalk", *Bioresource Technology*, 249, 574-581, 2017.

[30] Kumagai, S., Hirajima T. "Effective utilization of Moso-bamboo (Phyllostachys heterocycla) with hot-compressed water". In: Jin F, ed. *Application of Hydrothermal Reactions to Biomass Conversion, Green Chemistry and Sustainable Technology*. Berlin, Heidelberg, Springer, 155-170, 2014.

[31] El Khaldi-Hansen B, Schulze M, Kamm, B. "Qualitative and quantitative analysis of lignins from different sources and isolation methods for an application as a bio-based chemical resource and polymeric material". (Chapter 2). In: Vaz S Jr, ed. *Analytical Techniques and Methods for Biomass*. Switzerland, Springer, 15-44, 2016.

[32] Varol M., Atımtay A.T., Bay, B., Olgun, H. "Investigation of co-combustion characteristics of low quality lignite coals and biomass with thermogravimetric analysis", *Thermochim Acta*, 510, 195-201, 2010.

Numerical investigation of the effect of change of overlap tip geometry on strength in tubular adhesive joints with stress and failure analysis

İ.SARAÇ[1]

[1] Aksaray University, Aksaray/Turkey, ismailsarac@aksaray.edu.tr

Abstract - The use of adhesive joints is increasing due to the significant advantages it provides. One of the important advantages of adhesive joints is that it does not require any drilling in the parts to be joined and can easily combine different materials. It is possible to increase the strength of the bonding connections, especially with the design changes to be made in the bonding area. In this study, using ANSYS finite element software, pipe bonding joints models with different overlap end shapes were created. Then, stress and damage analyzes were made to investigate the effect of different overlap end shapes on the bond strength. The maximum stress theory defined in the ANSYS program was used to find the finite element failure loads of the modeled bonding bonds. As a result of the analyzes, significant strength increases were calculated according to the reference model. The strength increase in Mod-3(inside taper) and Mod-4 (inside recessing) was found to be 36% and 41 %, respectively, relative to the reference model.

Keywords **- Tubular adhesive joints; overlap tip geometry; stress analysis; failure analysis**

I. INTRODUCTION

IN addition to traditional fastening methods, the use of structural adhesives is increasing. Combining with adhesive has important advantages over other connection methods such as combining materials with different properties, homogeneous stress distribution in the bonding area and making the systems lighter.

Besides all these advantages, since there are many factors affecting the strength of adhesive joints, it is difficult to say anything about the bond strength in adhesive joints compared to conventional connection methods. Hence, adhesive joints are of interest to many researchers. Adhesive joints can be created in different design geometries. When the literature is examined, one of the joint types is tubular adhesive joints. The effects of many design parameters on the strength have been shown in the studies on tubular adhesive joints. Some of the basic design parameters investigated are overlap length, bonded material thickness, adhesive thickness and overlap tip geometry.

Pugno and Carpinteri investigated the stress distribution in the adhesive layer in pipe bonding joints under axial load. As a result of their stress analysis, the researchers determined that the maximum shear stress occurred at the edges of the bonding areas [1].

Nemeş et al. studied the effects of various parameters by constructing an analytical model in order to predict the stress distribution in cylindrical bonding joints exposed to tensile load [2].

Zou and Taheri developed an analytical model to examine the effects of factors such as material, pipe wall thickness, and adhesive type on cylindrical bonding joints with torsional loads [3].

Esmaeel and Taheri investigated the delamination effect in the cylindrical bonding joints they created using composite and aluminum materials using finite element method. As a result, they stated that in-layer separations affect the stress distribution in the adhesive layer [4].

Lyon developed a two-dimensional, symmetrical, finite element model that can be used in cylindrical bonding joints and other bonding joints in the linear elastic zone. As a result of the analysis, it has been shown that bonding zone length, thickness of bonded material, bonding zone edge geometries affect the bond strength [5].

Kumar and Khan studied the stress distribution in the adhesive layer in cylindrical bonding joints under tensile load by creating an analytical model. Then they made finite element analyzes in order to verify the analytical model they created. As a result of the study, they stated that analytical and finite element analysis are compatible [6].

Aimmanee et al. created an analytical model valid in the linear-elastic region in cylindrical bonding joints applied torsion moment. It is stated that the prepared model can be used for ideal connection conditions [7].

Noda et al. investigated analytically and numerically the damage behavior under tensile load in tubular bonding joints. In the study, it was stated that the region of the individual stress density is effective in the bond strengths [8].

Albiez et al. have experimentally investigated the effects of connection parameters such as pipe diameter, overlap zone length, pipe wall thickness, adhesive type, adhesive thickness, centering and angular non-linearity due to manufacturing to the strength under tensile load in cylindrical bonding joints. As a result of the study, the joint strength increased with the increase in the length of the overlap zone, but the increase was not linear, the increase in the adhesive thickness decreased the joint strength, the alignment error and angular non-linearities in the joined parts had less effect on the strength, the effect of the adhesive type and pipe wall thickness on the bond strength

was observed [9].

Lavalette et al. examined analytically and experimentally the effects of change in overlap zone length, inner pipe diameter change, pipe wall thickness, adhesive thickness, and overlap zone pipe end geometry on bond strength in aluminum-composite cylindrical bonding joints with tensile load. As a result of the parametric study, it was stated that the increase in the overlap length and the inner pipe diameter significantly affects the connection strength. The effect of overlap length increase on strength is not observed after a certain length. Changes in pipe wall thickness and pipe end geometry affected the stress distribution in the adhesive layer [10].

When the numerical analysis of tubular adhesive joints in the literature is examined, finite element models are usually created using the ANSYS program in two-dimensional axial symmetry. In this study, the effects of the overlap end geometry change on the stress distribution and failure in the adhesive region in tubular adhesive joints subjected to static tensile load were investigated by three dimensional finite element analysis. The three-dimensional models were created as a quarter model due to the symmetrical structure of the joint. The failure analysis was made according to the maximum stress theory included in the ANSYS program.

II. MATERIAL AND METHOD

Tubular adhesive joint models were created using ANSYS finite element software. The general dimensions and joint geometry of the models are shown in Figure 1. In all models created, the total length of the joint is 70 mm, the outer pipe diameter is 20 mm, the pipe wall thickness is 1 mm, and the adhesive thickness is 0.2 mm. The pipe material is steel and the adhesive is DP460 epoxy adhesive. Mechanical properties of steel and adhesive are shown in Table 1. Finite element analyzes were carried out as linear-elastic. In all analyzes performed, the elasticity module and poisson ratios of the steel material and adhesive given in Table 1 were defined in the ANSYS program. In order to estimate the finite element failure loads of the joints, the maximum stress failure criterion defined in the ANSYS program was used. According to this criterion, it is assumed that failure occurs when one of the stresses in the axes exceeds the allowable stresses. The limit values defined in the maximum stress failure criteria for pipes and adhesive materials have been selected for the yield strengths given in Table 1.

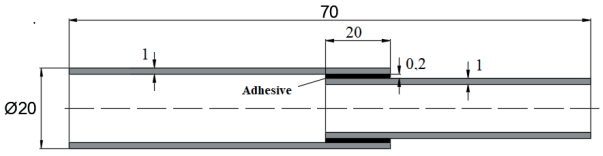


Figure 1: Tubular joint model overall dimensions

Table 1: Mechanical properties of DP460 and St37 [12,13].

Proporties	Materials	
	St37	DP460
Modulus of elasticity (MPa)	200000	1984
Poisson ratio	0,3	0,38
Yield strength (MPa)	235	38,4
Ultimate tensile strength (MPa)	375	44,6

Due to the cylindrical nature of the joint geometry, the model is symmetrical in the *yz* and *xz* planes. The models created considering this situation are designed as quarter model (1/4) (Figure 2). As a result of creating the models quarterly, it is aimed to reduce the number of element and to make the solution in a shorter time.

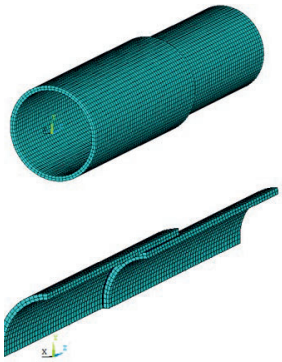


Figure 2: Full and quarter model

Meshing process is an important factor affecting the accuracy of the analysis. As a result of the convergence process, the element size was taken as 1 in all models. In the finite element models, solid 186 elements with 3 degrees of freedom and 20 nodes are used as shown in Figure 3.

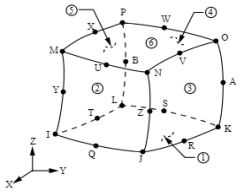


Figure 3: Solid 186 element [14]

The boundary conditions of the problem are defined in the program as shown in Figure 4. The pipe end with a large diameter is fixed in all directions. In all joints models, a tensile force of 3000 N is applied to the tip of the pipe with a small diameter. The symmetry boundary conditions for the *yz* and *xz* planes of the models are defined.

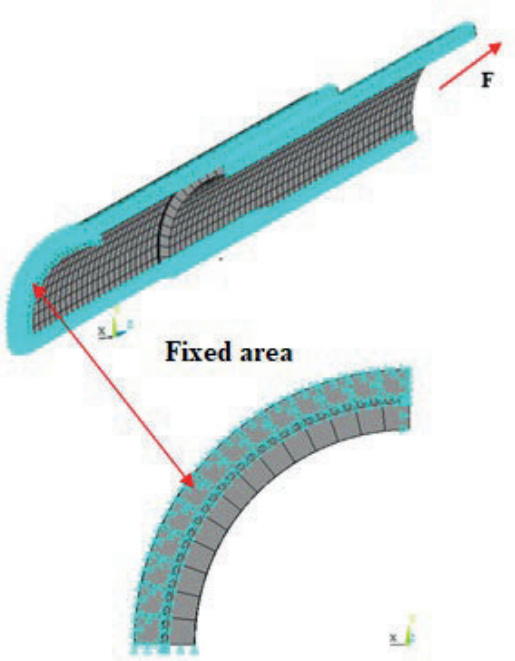


Figure 4: Boundary conditions of the joint models

In the study, the effects of the change of overlap tip geometry on the stress distribution in the critical area of the adhesive layer were examined. The stress distributions in the adhesive layer were obtained along the A-B line shown in Figure 5. The dimensions of the models created are shown in Figure 6.

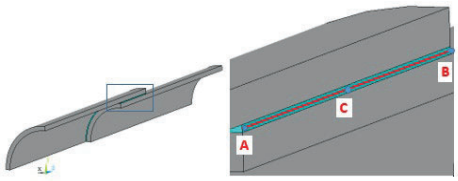


Figure 5: A-B line in adhesive layer

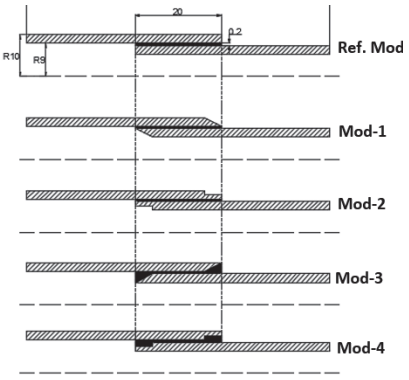


Figure 6: Models according to the overlap end geometry

III. Results and Discussion

In this study, the effects of the change of overlap end geometry on the stress distribution in the adhesive layer (A-B line) under tensile load at tubular adhesive joints were investigated by linear-elastic stress and failure analysis. The results obtained are shown in Figure 7-8. In the graphs, the points in the overlap region are normalized by dividing by the overlap length (X / L). For comparison, the force applied in all models is taken as 3000 N constantly. In the graphics shown in Figure 7-8, the normal stress (σ_y) and shear stress (τ_{yz}) distributions that are effective in the formation of failure in the adhesive layer are shown.

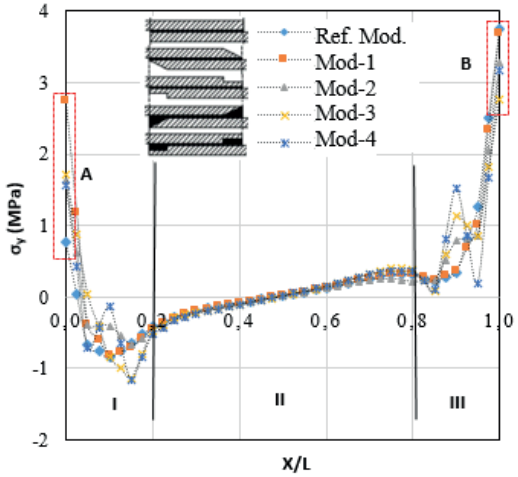


Figure 7: The effect of overlap tip geometry on σ_y stress distribution

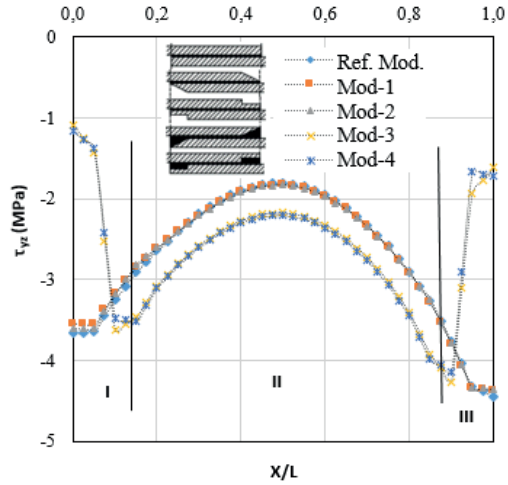


Figure 8: The effect of overlap tip geometry on the τ_{yz} stress distribution

Figure 7 shows the effect of overlap end geometry change on the σ_y normal stress distribution along the A-B line. In the tubular joint models, σ_y the normal stress values increase at the overlap end regions. When the σ_y normal stress distributions are examined along the A-B line, the stress

values with positive character at point A rapidly pass to the negative region and become zero at the C point (midpoint). It increases in the positive direction from point C to point B and takes maximum values at point B. As seen in Figure 7, σ_y normal stress values formed in the second region (II) are distributed more evenly compared to the first (I) and third (III) regions.

When the effects of the changes in overlap end geometry on the σ_y normal stress distribution along the A-B line is examined, it is seen that the effects occurring in the first and third regions, especially at the A and B points, are different. At point A, the minimum stress occurred in the reference model. At point B, the minimum stress occurred in the Mod-3 design (inside taper). Point B is the point closer to the end where the tensile load is applied. At the same time, it is the part where the tension is transferred from the inner pipe to the outer pipe.

Figure 8 shows the effect of overlap tip geometry change on the τ_{yz} shear stress distribution along the A-B line. When the effect of the overlap geometry changes on the τ_{yz} shear stress distribution along the A-B line is examined, Ref. Mod., Mod-1 and Mod-2 designs and Mod-3, Mod-4 designs differ from each other by showing a significant difference especially in the overlap end zone (I and III. zone). In Ref. Mod., Mod-1 and Mode-2 designs, the maximum τ_{yz} shear stress values occur at the A and B points. In Ref. the Mod., Mode-1 and Mode-2 designs, τ_{yz} shear stress distributions are quite close to each other along the A-B line. In Mod-3 and Mode-4 designs, τ_{yz} shear stress values occurring at A and B points are at minimum values and τ_{yz} shear stress distributions formed along the A-B line coincide. The maximum τ_{yz} shear stress values in Mod-3 and Mod-4 designs were formed approximately 1 mm inside the A and B points.

Finite element failure loads of the joints are shown in Table 2. Failure loads were calculated using the maximum stress failure index coefficients obtained according to the failure criterion determined in the ANSYS program. Failure loads were determined by dividing the amount of load applied to the joint by the failure index coefficient. According to the results, the finite element failure load increased in all models compared to the reference model. Significant failure load growth rates were calculated as 36% and 41% in Mod-3 and Mode-4, respectively, compared to the reference model. As a result, with the decrease in the stress concentration at the overlap ends, failure loads increase.

Table 2: Finite element failure loads of the joints

Model type	Applied load (N)	Failure index	Finite element failure load (N)
Ref. Mod.	3000	0,242578	12367
Mod-1	3000	0,238356	12586
Mod-2	3000	0,233387	12854
Mod-3	3000	0,177896	16863
Mod-4	3000	0,172205	17421

IV. Conclusion

In this numerical study, the effects of the change of overlap tip geometry on the stress distribution and failure in tubular adhesive joints under tensile load have been examined by three dimensional finite element analysis. General results obtained are given below.

Failure loads increased with the decrease in the stress concentration at the overlap ends in the created models. Significant failure load growth rate was calculated as 36% and 41% in Mod-3 and Mode-4, respectively, compared to the reference model.

Overlap tip geometry change affected the stress distribution in the adhesive layer. The stress distribution characteristics seen in the Ref. Mod., Mode-1 and Mode-2 designs are similar. Likewise, the stress distribution characteristics of the Mod-3 and Mode-4 designs are similar.

The normal stress values of σ_y are higher at the end of the overlap which is close to the force applied part, compared to the other overlap end in all models.

When the τ_{yz} shear stress distribution in the overlap region is examined, In the Ref. Mod., Mod-1 and Mod-2 designs, maximum stresses occurred at the overlap ends. In Mod-3 and Mod-4 designs, the maximum stresses were formed at the inner parts of the overlap ends 1 mm.

References

[1] V Nguyen, T Kedward , "Non-linear modeling of tubular adhesive scarf joints loaded in tension," *The Journal of Adhesion*, 76(3), pp 265-295, 2001

[2] N Pungo, A Carpinteri, "Tubular adhesive joints under axial load," *Journal of Applied Mechanics*, 76(6), pp 832-839, 2003

[3] O Nemes, F Lachaud, A Mojtabi, "Contribution to the study of cylindrical adhesive joining," *International Journal of Adhesion & Adhesives*, 26, pp 474-480, 2006

[4] G P Zou, F Taheri, "Stress analysis of adhesively bonded sandwich pipe joints subjected to torsional loading," *International Journal of Solids and Structures,* 43(20), pp 5953-5968, 2006

[5] R A Esmaeel, F Taheri, "Stress analysis of tubular adhesive joints with delaminated adherend," *Journal of Adhesion Science and Technology,* 23, pp 1827-1844, 2009

[6] P E Lyon, "Axisymmetric finite element modeling for the design and analysis of cylindrical adhesive joints based on dimensional stability," Master of Science*,* Utah State University, Logan, Utah, 2010

[7] S Kumar, M A Khan, "An elastic solution for adhesive stresses in multi-material cylindrical joints," *International Journal of Adhesion & Adhesives*, 64, pp 142-152, 2016

[8] S Aimmanee, P Hongpimolmas, "Stress analysis of adhesive-bonded tubular-coupler joints with optimum variable-stiffness composite adherend under torsion," *Composite Structures,* 164, pp 76-89, 2017

[9] N A Noda, F Ren, R Takaki, Z Wang, K Oda, T Miyazaki, Y Sano "Intensity of singular stress field over the entire bond line thickness range useful for evaluating the adhesive strength for plate and cylinder butt joints," *International Journal of Adhesion and Adhesives,* 85, pp 234-250, 2018

[10] M Albiez, T Vallée, H Fricke, T Ummenhofer, "Adhesively bonded steel tubes Part I: Experimental investigations," *International Journal of Adhesion and Adhesives,* 90, pp 199-210, 2019

[11] N P Lavalette, O K Bergsma, D Zarouchas, R Benedictus, "Influence of geometrical parameters on the strength of Hybrid CFRP-aluminium tubular adhesive joints," *Composite Structures,* 240, pp 112077, 2020

[12] S Akpınar, "The strength of the adhesively bonded step-lap joints for different step numbers," *Composites Part B,* 67, pp 170-178, 2014

[13] L Malag, L Kukielka, "Hybrid method to determinate the states of deformation and stress in material during the tensile test," Proc *Appl. Math. Mech.,* 7, pp 2090025–2090026, 2007

[14] https://ansyshelp ansys com

Free Vibration Analysis of Asymmetric Sandwich Structure with ANSYS

Ufuk DEMİRCİOĞLU[1] and M. Tarık ÇAKIR[1]

[1] Sivas Science and Technology University, Sivas/Turkey, udemircioglu@sivas.edu.tr
[1] Sivas Science and Technology University, Sivas/Turkey, tcakir@sivas.edu.tr

Abstract - **Asymmetric sandwich structure is a new construction method in which face sheet is no longer made from same materials. Thanks to this configuration, it gives more choice to designer to obtain desired properties in sandwich structures by changing or mixing the face sheet materials. This study presents the free vibration analysis of asymmetric sandwich structures comparatively. Sandwich structures were constructed and solved by using commercially available software ANSYS. Free vibration behavior of asymmetric sandwich structure was evaluated by comparing it with symmetric sandwich structures. Finite element method was used to analysis the free vibration analysis of sandwich structure. The first six modes of sandwich structures were obtained. It was obtained from analysis that as the face sheet changes the natural frequencies also change. The analysis showed that asymmetry has effect on natural frequencies of sandwich structure**.

Keywords **- Sandwich Structures, Low Density Sandwich Structures, Asymmetric Sandwich Structures, Vibration Analysis**

I. Introduction

Low density sandwich structures are nonhomogeneous sub group of composite materials [1]. Sandwich structures are manufactured from two thin face sheets and one thick low-density core material which is adhered between face sheets [2]. Generally, face sheets materials carry tensile and compressive loads. There are a numerous of face sheets that includes carbon fiber, glass fiber and any thin sheet metals. The objective of usage of the low-density core materials in sandwich structure is to increase the sandwich structure's thickness and as a result to increase stiffness of sandwich structure [3]. This gives a dramatic increase in stiffness for very little addition of weight. There are various core materials such as polyurethane foams, Poli Vinil Clorur (PVC) foams, honeycomb cores, wood cores (balsa and cedar). As a result, sandwich structures have high strength-to-weight and stiffness-to-weight ratios and therefor they are used as light weight load bearing components where needed [4]. Sandwich construction offers better stiffness, energy absorption, and shock resistance than that of conventional metallic materials. As a result, Sandwich structures are widely used in aerospace, automotive, defense industry due to their versatile properties.

Kumar et al. [5] studied transient response, harmonic response and static response of the sandwich beam subject to a concentrated point load at the mid span of the beam. The effect of the core thickness on vibration is evaluated. Finite element method is used by using ANSYS to analyze the overall transient responses, harmonic responses and the static responses of the sandwich structure. A simple software has been developed for calculation of data and the results are compared with ANSYS results to validate. Purushothaman and Guruprasad [6] simulated the natural frequencies and mode shapes of the sandwich beam structure under different core configuration with different core materials buy using ANSYS. From results, it was obtained that truss core is better which has higher stiffness and circular core responds effectively. Şakar et al. [7] studied vibration behavior of aluminum honeycomb sandwich structures experimentally and numerically with ANSYS. The natural frequencies and mode shapes of sandwich structures fabricated with different configurations for clamped-free and free-free boundary conditions were determined. The effects of lower and upper face sheet thickness, the core material thickness and foil thickness on the vibration characteristics were examined. Marythraza et al. [8] studied the normal mode

and transient analysis of honeycomb sandwich structures made up of aluminum honeycomb core and aluminum face sheet by using Msc Nastran/Patran for spacecraft application. Simulation results were compared with analytical results. From analysis it is obtained that transient respond of the panel fills the requirement. AL-Araji et al. [9] studied free and forced vibration of perforated sandwich plate. the natural frequency and mode shape of the sandwich panels have been obtained by using finite element analysis with ANSYS. The effect of number of holes and perforation ratio on vibration was investigated. Ersin et al. [10] investigated the free vibration of sandwich beam made of functionally graded materials method. Natural frequencies were computed for simply-supported boundary conditions of beam by using Timoshenko and Euler-Bernoulli theorems. The effect of slenderness ratio and power index are discussed. The obtained results are compared with the ANSYS result.

All of the studies have considered symmetric sandwich structure when analyzing. However, asymmetric sandwich structures have been used and investigated for nearly 20 years. Asymmetric sandwich is a new type of sandwich structure configuration with more properties in that face materials, geometrical size and material properties different from each other [11]. This configuration can offer more properties to designers than that of symmetric configuration. Therefor asymmetric sandwich structures have been used widely especially in defense industry [12],[13],[14]. so, it is of the importance to study free vibration behavior of asymmetric sandwich structure.

In this study, free vibration analysis of asymmetric sandwich structures will be evaluated comparatively by using commercially available software ANSYS. A simple sandwich structure is shown in figure 1.

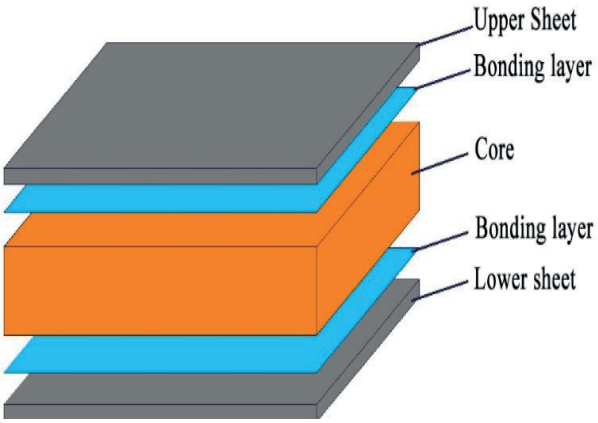


Figure 1: Sandwich structure

II. MATERIALS AND METHOD

Sandwich structures are modeled by using ANSYS workbench commercially available software. Three different sandwich structures namely, two of them symmetric and one of them asymmetric are created by using ANSYS ACP tool (Table 1.). Symmetric sandwich structures have carbon fiber and glass fiber on both face sheets respectively. Asymmetric sandwich structure has carbon fiber on one face and glass fiber on the other face. In all the sandwich structure same PVC foam is used. Sandwich structure that has carbon fiber on both faces, glass fiber on both faces and carbon fiber on one face and glass fiber on the other face is denoted as C-C, G-G and C-G respectively (Table 1.). Lay-up of sandwich structures is given in table 2. Dimensions of sandwich structure and thickness of carbon fiber, glass fiber and PVC foam are given in figure 2, table 3 and table 4 respectively.

Table 1: Configuration of Sandwich structures.

C-C	G-G	C-G
Carbon Fiber	Glass Fiber	Carbon Fiber
PVC Foam	PVC Foam	PVC Foam
Carbon Fiber	Glass Fiber	Glass Fiber

Table 2: Lay-up of sandwich structures.

90
0
90
0
Foam
0
90
0
90

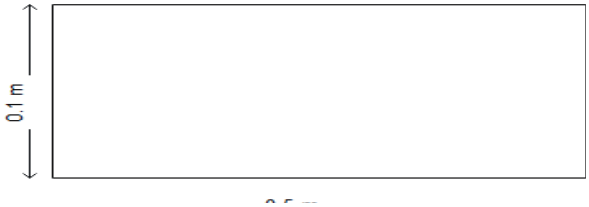


Fig 2: Dimensions of sandwich structure.

Table 4: Thickness of the materials.

Thickness of carbon fiber	0,0002 m
Thickness of glass fiber	0,0002 m
Thickness of PVC foam	0,001 m

Materials from ANSYS library is used in construction of sandwich structures. Properties of PVC foam is given in table 5,6 similarly, properties of carbon fiber and glass fiber is given in table 7,8 respectively.

Table 5: Density and Isotropic Elasticity of PVC foam.

Density kg m^-3	Young's Modulus Pa	Poisson's Ratio	Bulk Modulus Pa	Shear Modulus Pa
60	7x10^7	0,3	5,8333 x10^7	2,6923 x10^7

Table 6: Orthotropic Stress Limits.

Tensile X direction Pa	Tensile Y direction Pa	Tensile Z direction Pa	Compressive X direction Pa	Compressive Y direction Pa
1,5x10^6		9,5x10^6	-1,5 x10^6	-1,5 x10^6
Compressive Z direction Pa	Shear XY Pa	Shear YZ Pa	Shear XZ Pa	
-9,5 x10^6	9,3x10^5	9,3x10^5	9,3x10^5	

Table 7: Density and Isotropic Elasticity of glass fiber

Density kg m^-3	Young's Modulus Pa	Poisson's Ratio	Bulk Modulus Pa	Shear Modulus Pa
2600,	7,3 x10^10	0,22	4,3452 x10^10	2,9918 x10^10

Table 8: Orthotropic Stress Limits.

Density kg m^-3	Young's Modulus X direction Pa	Young's Modulus Y direction Pa	Young's Modulus Z direction Pa	Poisson's Ratio XY
1800,	2,3x10^11	2,3x10^10	2,3x10^10	0,2
Poisson's Ratio YZ	Poisson's Ratio XZ	Shear Modulus XY Pa	Shear Modulus YZ Pa	Shear Modulus XZ Pa
0,4	0,2	9 x10^9	8,2143x10^9	9 x10^9

III. RESULTS AND DISSCUSSION

Model analysis of sandwich structure is performed in ANSYS by using finite element method. Analysis is performed when one end of sandwich structure is fixed while the other end is free. First six frequencies of each sandwich structures are obtained from results as shown in table 9. Mode shapes of asymmetric sandwich structure is shown in figure 3, 4, 5, 6,7 and 8. When symmetric structures are considered, namely, when carbon fiber is used as face sheet material on both faces, natural frequencies are the highest, when glass fiber is used as face sheet material on both faces natural frequencies are the lowest. When asymmetric structure is considered, that is to say, when glass fiber and carbon fiber are used together as face sheet materials natural frequencies are different than symmetric sandwich structures. From results one can obtain that asymmetry has effects on natural frequencies of sandwich structures.

Table 9: Frequencies of sandwich structures.

Mode	Frequency [Hz] C-C	Frequency [Hz] G-G	Frequency [Hz] C-G
1.	78,14	54,779	64,08
2.	224,43	239,32	224,71
3.	311,46	243,16	269,67
4.	438,9	312,3	376,18
5.	641,75	520,36	566,13
6.	660,34	700,24	658,38

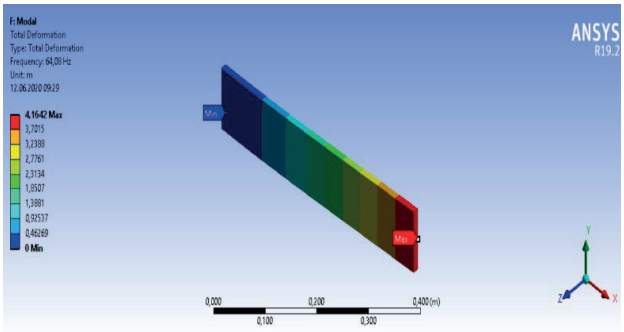


Fig 3: First mode of asymmetric sandwich structure.

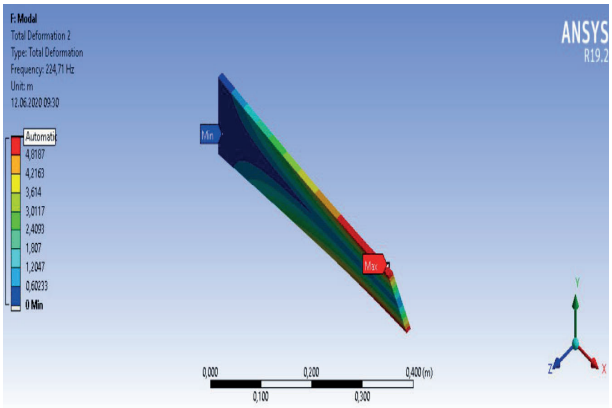


Fig 4: Second mode of asymmetric sandwich structure.

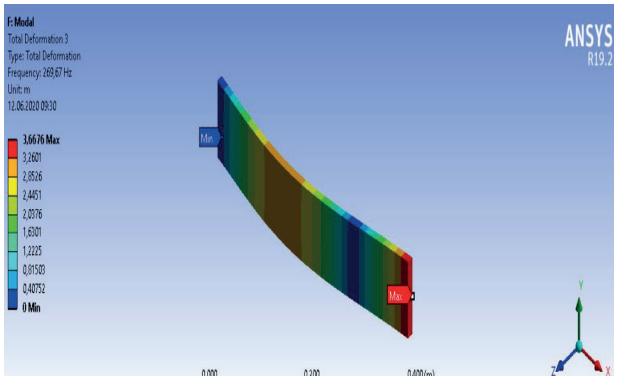


Fig 5: Third mode of asymmetric sandwich structure.

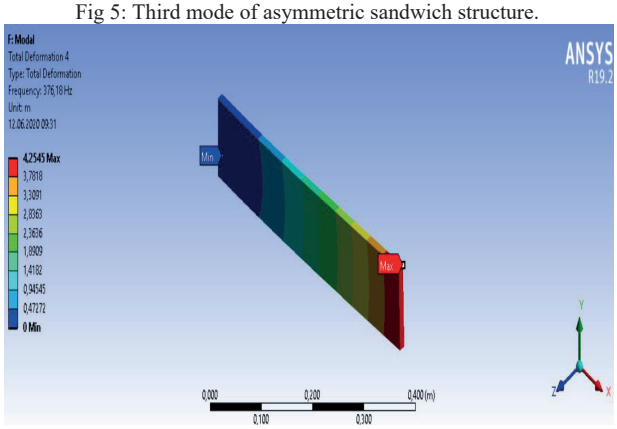


Fig 6: Fourth mode of asymmetric sandwich structure.

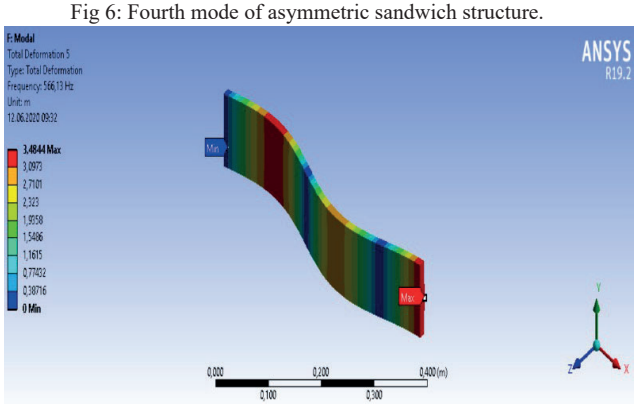


Fig 7: Fifth mode of asymmetric sandwich structure.

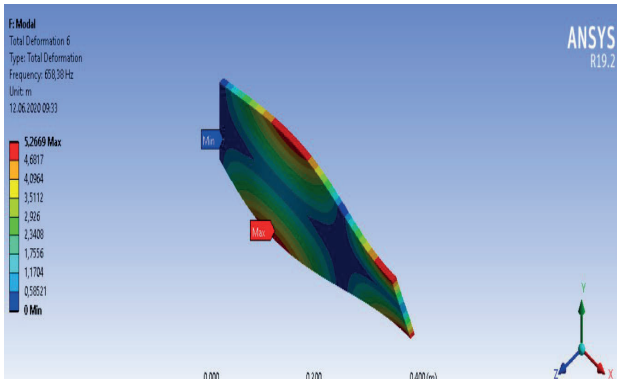


Fig 8: Sixth mode of asymmetric sandwich structure.

IV. CONCLUSION

In this study, vibration behavior of asymmetric sandwich structure was evaluated in ANSYS by using finite element method. Analysis were performed when sandwich structures are fixed at one end and free at the other end. Results obtained from analyzing sandwich structures were investigated comparatively. From the analysis, it is obtained that face sheet materials have effect on vibration behavior of sandwich structures. Based on the result one can see that as the face material shift from one to another natural frequencies of sandwich structures change. This gives to designer the opportunity to keep the frequency at desired level by manipulating face sheet materials.

REFERENCES

[1] A Krzyzak, M Mazur, M Gajewski, K Drozd, A Komorek, and P Przybyłek, "Sandwich Structured Composites for Aeronautics: Methods of Manufacturing Affecting Some Mechanical Properties," *Int. J. Aerosp. Eng.*, vol 2016, 2016, doi: 10 1155/2016/7816912

[2] M Rao and S P Mitra, "Static and Vibration Analysis of Composite Beams Using Finite Element Methods," *Int. J. Res.*, vol 2, no 11, pp 75–80, 2015

[3] O Mohammed, E Suleiman, A F A Algarray, and I M Mahdi, "Free Vibration Analysis of Composite Laminated Beams International Journal of Engineering Research And Free Vibration Analysis of Composite Laminated Beams," *Int. J. Eng. Res. Adv. Technol.*, no July, 2019

[4] J Xiong, Y Du, D Mousanezhad, M Eydani Asl, J Norato, and A Vaziri, "Sandwich Structures with Prismatic and Foam Cores: A Review," *Adv. Eng. Mater.*, vol 21, no 1, pp 1–19, 2019, doi: 10 1002/adem 201800036

[5] E V Prakash, M Kumar, K Sinha, and A Shilpi, "To study analysis of sandwich beam with ansys and fem," *Int. J. Inf. Res. Rev.*, vol 2, no April, pp 947–951, 2015, [Online] Available: https://www researchgate net/publication/301557907

[6] V Purushothaman, "Vibration Analysis of Sandwich Beam with Different Core Patterns," *Int. J. Eng. Res. Technol.*, vol 6, no 06, pp 2–5, 2018

[7] G Şakar, M Yaman and F C Bolat, "Alüminyum Bal Peteği Sandviç Yapıların Serbest Titreşim Analizi", 5 Otomotiv Teknolojileri Kongresi, March, 2018

[8] M Marythraza, D Anitha, P K Dash, and P Ravi Kumar, "Vibration analysis of honeycomb sandwich panel in spacecraft structure," *Int. J. Mech. Prod. Eng. Res. Dev.*, vol 8, no 3, pp 849–860, 2018, doi: 10 24247/ijmperdjun201890

[9] M S H Al-araji and A S Gafer, "Vibration Analysis of Perforated Composite Sandwich Plate," no 26, pp 166–178, 2018

[10] E Demir, H Çallioğlu, and M Sayer, "Free vibration of symmetric FG sandwich Timoshenko beam with simply supported edges," *Indian J. Eng. Mater. Sci.*, vol 20, no 6, pp 515–521, 2013

[11] C Yuan, Q Qin, and T J Wang, "A Yield Criterion for Hybrid Asymmetric Metal Sandwich Structures and its Application," *Int. J. Appl. Mech.*, 2018, doi: 10 1142/S1758825116500010

[12] A A Khurram, S A Rakha, N Ali, M T Asim, Z Guorui, and A Munir, "Microwave Absorbing Properties of Lightweight Nanocomposite/Honeycomb Sandwich Structures," *J. Nanotechnol. Eng. Med.*, vol 6, no 1, pp 1–6, 2015, doi: 10 1115/1 4031472

[13] A A Khurram, M A Raza, P Zhou, and T Subhani, "A study of the nanocomposite sandwich structures for broadband microwave absorption and flexural strength," *J. Sandw. Struct. Mater.*, vol 18, no 6, pp 739–753, 2016, doi: 10 1177/1099636216650837

[14] M A M Hunjra, M A Fakhar, K Naveed, and T Subhani, "Polyurethane foam-based radar absorbing sandwich structures to evade detection," *J. Sandw. Struct. Mater.*, vol 19, no 6, pp 647–658, 2017, doi: 10 1177/1099636216635856

Stochastic Optimization of Process Parameters for Pure Titanium Using Wire Electric Discharge Machining (WEDM) Process

Y.M.BAYTOK[1], L. AYDIN[2], E. GÜLTÜRK[3]

[1] İzmir Katip Çelebi University, İzmir/Turkey, ymbaytok@gmail.com
[2] İzmir Katip Çelebi University, İzmir/Turkey, leventaydinn@gmail.com
[1]Kocaeli University, Kocaeli/Turkey, elifgulturk@gmail.com

***Abstract* - This study aimed to model and find the most suitable solutions for cutting parameters during wire electrical discharge machining (WEDM) so that pure titanium plate (Grade-2) has a smooth cutting process. In this study, the effects of cutting parameters such as peak current (IP), pulse off time (T_{off}), pulse on time (T_{on}), spark gap voltage (SV), wire feeding (WF), and wire voltage (WT) on surface roughness (SR) are investigated using Wolfram Mathematica. Different multiple nonlinear regression models were used to define the cut parameters of the WEDM. The coefficient of determinations (R^2) and the adjusted coefficient of determinations ($R^2_{adjusted}$) of each model was calculated to see how well models identify the cutting parameters. Besides, the stability of the models was determined to measure the ability of the WEDM to realize the cutting parameters. The stability results showed that the selected model was suitable for the phenomena. After the stability calculations are done optimum values of surface roughness are determined by Simulated Annealing. By using multiple nonlinear regression analysis and optimization process, the optimum value of surface roughness (SR) was obtained as 2.06µm. This study also shows the importance of mathematical modeling of engineering problems for real-life applications.**

***Keywords* - Stochastic Optimization, Multiple Nonlinear Regression Analysis, Wire EDM, Stability, Surface Roughness**

I. INTRODUCTION

TITANIUM has numerous usage areas in critical applications, mostly in industry, due to its fracture resistance, resistance to corrosion in salty and acidic environments, and its high strength value per weight at high temperatures. These are the aerospace industry, petroleum refinery, airframe construction, nuclear reactors, marine applications, and surgical implants. Additionally, it is used in compressor blades, offshore pressure vessels, and rocket casings [1, 2]. Machining of titanium and its alloys is very difficult due to its tendency to weld with cutting tools during machining and its superior strength at high temperatures. Titanium has low thermal conductivity, which causes an increase in temperature between the tool and the workpiece during machining. This situation causes shortening of tool life and earlier breakdown [3]. Due to the properties of titanium, it is very difficult to machine with traditional machining techniques such as turning, milling, etc. Recently, nonconventional machining operations such as laser beam machining (LBM) and electric discharge machining (EDM) have been used for the machining of pure titanium and its alloys, however, some limitations should be taken into consideration. The most important of these are surface finish, dimensional inaccuracies as well as undesirable conditions such as heat-affected zone, thermal stresses, and recast layer. One problem with EDM processing is that residual material in the processing cavity cannot be easily washed and an unstable processing state occurs [4]. Re-casting layer means that some materials melt and re-solidification happens on the base materials, resulting in a different surface coating than the base material [1].

Wire-EDM is a recently introduced electro-thermal metal removal process that is widely preferred in nuclear and aerospace industries to produce complex shapes with tight tolerances and fine surface quality [5]. WEDM is an erosion process in which a continuously moving wire acts as a cutting tool. Material removal in WEDM is mainly related to the effect of erosion due to spark discharges between the wire and the workpiece material being immersed in a dielectric media. In WEDM, no direct contact occurs between the workpiece and wire, and; as a consequence, the undesirable effects that normally present in conventional machining such as mechanical stresses, chatter, and vibration are diminished [6]. The Wire-EDM process has various cutting parameters such as, Pulse current, Pulse-on time, pulse-off time, applied voltage wire-feed rate, wire tension, spark gap voltage, injection pressure which has an influence on its performance [7]. Since these parameters affect properties such as surface roughness, material removal rate, dimensional deviation, and wire wear ratio, it is important to choose the optimum combination of WEDM parameters to ensure that the machined parts can pass the quality control and ensure high surface integrity at high production rate. Scott et al. [8] tried to optimize cutting parameters of WEDM and found that whenever pulse duration, discharge current, and wire-speed increase, the surface of specimens is getting better. On the other hand, they showed that the surface finish is getting worse by reducing the dielectric flow rate. Spedding and Wang [9] were used an ANN model to optimize the effect of process parameters on Inconel 718 material surfaces machined with WEDM. Finally, it is optimized the cutting rate and

surface roughness values. Tosun et al. [10] attempted to find the relationship between the diameter, depth of erosion craters on the wire, and cutting parameters. The higher dielectric fluid flushing pressure leads to reducing the crater size. On the other hand, increasing the pulse duration, open-circuit voltage, and wire-speed cause smaller craters to form. In another study, the effects of cutting characteristics on surface roughness were examined based on experimentations in the WEDM process. The study revealed that surface roughness improved with the reduction of discharge current and pulse duration. Constant impact energy/discharge ratio means that short and long pulses occur the similar surface roughness. However, surface morphology and material removal rates will be different. The removal rate rises with a shorter pulse direction [11].

The aim of this study is to model and finding optimal solutions for cut parameters during wire electrical discharge machining (WEDM) of pure titanium plate (Grade-2). The effects of six cutting parameters, pulse on time (T_{on}), pulse off time (T_{off}), peak current (IP), spark gap voltage (SV), wire feeding (WF), and wire voltage (WT) on surface roughness (SR) were investigated by using Mathematica. As a result, the Best fitness values of cutting parameters of a titanium plate in WEDM and the corresponding optimal values are calculated by using Simulated Annealing optimization.

II. MATERIALS AND METHOD

A. *Design of Experiment*

Experimental matrices in the study provided by Box-Behnken design. Based on pilot experiments, “one variable at a time” was used to determine the factors and levels. The experiments were applied to six different cutting parameters such as pulse on time, pulse off time, peak current, spark gap voltage, wire feed, and wire tension with different values , and the effects of these parameters on surface roughness were investigated. Pulse on time (T_{on}), pulse off time (T_{off}), peak current, and spark gap voltage setting (SV) are the most significant parameters [12]. In order to improve the WEDM processes of pure titanium and to obtain a smooth cut, the input parameter level was used to optimize the parameters [1].

Table 1: Factors and their levels [1]

S.NO	Input factors	Level			Units
		1	2	3	
1	T_{on}	112	116	120	µs
2	T_{off}	44	50	56	
3	I_p	120	160	200	Ampere
4	SV	40	50	60	Volt
5	WF	4	7	10	m/min
6	WT	500	950	1400	

Box-Behnken design, which consists of constructing incomplete and balanced block scans, is explained by three factors. While designing in this method, one factor is kept at a constant level and combinations are applied at all levels for other factors [13]. In this study, the third-factor level was fixed and all combinations for all levels of the other two factors were formed. Then, the same operations were applied for factors II and I. center point values are located in the last column of the created design matrix [14]. 54 experiments taken from an article [1] were based on the Box-Behnken design method to find the values of independent factors The independent factors and levels of these factors are given in Table 1.

B. *Experimental Background and Data Collection*

A four-axis CNC type WEDM (Electronica Sprintcut 734) given schematically in Figure 1 was used to implement the experiments. A set of experiments were carried out by changing a total of six cutting parameters consisting of pulse on time, pulse off time, peak current, spark gap voltage, wire feed, and wire tension. In all experiments, the material was pure titanium, the diameter of the wire-electrode was 0.25mm, the material thickness was 26 mm and the dielectric pressure was 7 kg/cm^2. Surface roughness was measured three times for each sample using Micrometer (Mitutoyo's SURFTEST (SJ-301) and the average roughness was determined. Square plates of 26 mm thickness and 148 mm × 148 dimensions were prepared for the experiment. The measured machining rate was taken as mm / min.

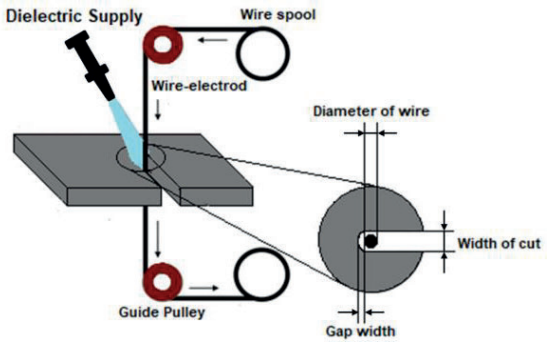


Figure 1: Schematic representation of WEDM [15].

Fig. 2 presents a profile showing the work-path throughout machining. The point of O (5, 0) was decided to be the reference. The path of the wire is determined as OA-AB-BC-CD-DO. All experiments, the unit of time and pulse interruption time in microseconds (µs). For the wire tension, measurements were made in grams. During machining, the wire offset was set at zero. ELAPT software was used to compose the CNC program [1].

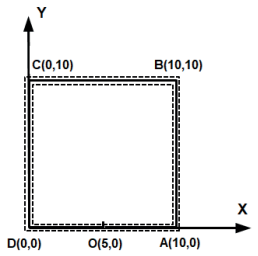


Figure 2: Profile showing Work-Path [1].

C. Regression Analysis

It is a robust and reliable method that allows you to model an "output" as a function of "design variables". Regression analysis is utilized in the prediction of the parameters' relationship which affects the engineering processes. The most widely preferred regression forms are Linear, Logistic, Nonlinear, Rational, and Stepwise in the literature. Whether the regression models are good or not is determined by R^2, and it is desired to be in the 0-1 range and close to 1. Moreover, another model-data compatibility measure, $R^2_{adjusted}$, decreases when unnecessary terms are added to the model.

Since only the determination of R^2 is not enough to define the engineering process/phenomenon, $R^2_{adjusted}$ is also calculated to measure how close the fitted model results to the experimental data. As a result, the less the difference between R^2 and $R^2_{adjusted}$, the better the model can reflect the phenomenon. It is understood from the Eq. (4) and $R^2_{adjusted}$ will be smaller or equal toR^2. The difference between them depends on the number of experiments for the same independent design variables and if the number of data is increased, the difference between them will decrease, so it is important to get accurate results [16]. These parameters can be formulated as in the following form:

$$R^2 = 1 - \frac{SSE}{SST} \quad (1)$$

Eq. (1) contains sum of square errors (SSE) and total sum of squares (SST) which are determined before calculation of R^2 Equation 2 and 3 are represents the formulation of them.

$$SSE = \sum_{i=1}^{n}(f_i - f_i')^2 \quad (2)$$

$$SST = \sum_{i=1}^{n}(f_i - \bar{f})^2 \quad (3)$$

f_i is the fitted values, f_i' is the corresponding point on the regression line, and $\bar{f}$ represents *the mean* f -value for the dataset.

$$R^2_{adjusted} = 1 - (1 - R^2)\frac{(n-1)}{(n-k-1)} \quad (4)$$

Multiple Nonlinear Regression

Although multiple nonlinear regression is similar to simple linear regression, it differs in that the model functions are nonlinear and the input number is more than one [17].

The fitted model can be used to estimate the SR value with the new observed values of T_{on}, T_{off}, IP, SV, WF, WT.

The First order trigonometric non-linear regression model used in this article are given in Table 2.

Table 2: The proposed regression model for the present application

Model Name	Nomenclature	Formula
First Order Trigonometric Non-linear	FOTN	$a_0 + a_1 sinx1 + a_2 \sin x2 + a_3 sinx3 + a_4 sinx4 + a_5 \sin x5 + a_6 sinx6 + a_7 cosx1 + a_8 \cos x2 + a_9 cosx3 + a_{10} cosx4 + a_{11} cos\ x5 + a_{12} cosx6$

Regression Analysis In Mathematica

In this study, the "***FindFit*** " solver is used to obtain coefficients of multiple nonlinear regression models in Mathematica. ''ConjugateGradient'',''Gradient'', LevenbergMarquardt'', ''Newton'', and ''QuasiNewton'' are the method embedded in the solver during calculation of coefficients.

D. Neuro Regression Approach

In the neuro regression approach, the idea of artificial neural networks is used. Experimental data is separated as 20 % and 80 % in a random way. 80 % of all data was used for training. During the neuro regression step, the "RandomSample" command is used to take 80 % of data from the dataset. After that, the "Complement" command takes the remaining data from the dataset. By using training data, a training model was created. 20% of all data was used as the testing data. After the process the predicted training model matched well with WEDM reactions, hence, resultingR^2, $R^2_{adjusted}$, $R^2_{training}$and $R^2_{testing}$ values were higher than 0.85.

Stability

The stability of the model is determined in order to measure the ability to realize the WEDM parameters. Stability results are represented in order to ensure if the chosen model is suitable or not. In this study, the behavior of the models was investigated. Mathematical models of one output parameters have been increased and decreased by %10, %30, and %20. By this means; it is seen at which intervals the models were stable. Stability calculations have been done for the objective function of SR.

E. Optimization method (Simulated Annealing (SA))

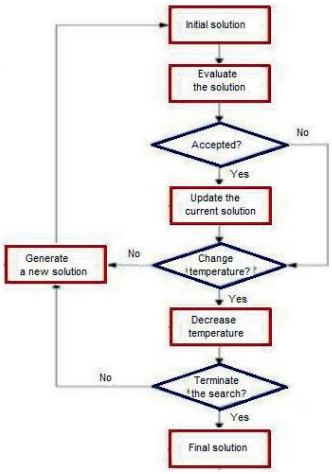


Figure 3: Flowchart of the SA Method [Adapted from [18]]

Mathematica has been used in this study to perform the stochastic optimization process. In this process, ***NMinimize*** and ***NMaximize*** solvers are selected to obtain optimal values of cutting parameters of Wire EDM. Based on the deterministic nature of variables, the objective function was minimized considering the system constraints and bounds by

Simulated Annealing optimization algorithms The flowchart of the SA model is given Fig.3.

SA is an analogy between the annealing process and optimization. When steel annealing is taken into consideration, it is a heat treatment method that is used to cool the steel from high temperatures to low temperatures, slowly. The main idea of annealing is to improve machinability, ductility, and decreasing the internal stresses [19, 20].

F. Problem Definition

The optimization process was carried out in Mathematica in order to obtain maximum efficiency from the performance of the WEDM machining. All experimental data were taken from the article of Kumar et.al [1]. This study aims to define optimum values of important cutting parameters such as Pulse off time (T_{off}), Pulse on time (T_{on}), Peak current (IP), Spark gap Voltage (SV), Wire Feed (WF), and Wire Tension (WT) in Wire EDM.

In this paper, the first purpose is to create a non-linear regression model, by using the experimental data which are taken from another study on the same problem to search the high-quality optimal process parameter conditions. The goal was to minimize the Surface Roughness.

To obtain the best result for the response parameters these steps must be followed.

Define the Objective Function,
Design Variables,
Constraints
Modeling

III. RESULTS AND DISCUSSION

A. Modeling of SR

At the end of the several trials to choose an appropriate model for SR which represents phenomena the best, a decision was made on the first-order multiple nonlinear trigonometric regression model (FOTN). The reason behind why the FOTN model is used for the process parameters of SR is the good relation between response and actual data. At the neuro regression step, after training and testing dataset created, $R^2_{training}$ and $R^2_{testing}$ values of the proposed model were found and given in Table 3.

Table 3: Results of R^2, $R^2_{adjusted}$, $R^2_{training}$ and $R^2_{testing}$ values.

Output	R^2	$R^2_{adjusted}$	$R^2_{training}$	$R^2_{testing}$
SR	0.95	0.89	0.95	0.88
SR= 3 65152 -0 0116287 Cos[Ip]+0 0441505 Cos[SV]-0 791602 Cos[Toff]+0 0838504 Cos[Ton]-0 00539295Cos[WF]+ 0 0230334 Cos[WT]-0 103772 Sin[Ip]+0 163232Sin[SV]+ 0 899594Sin[Toff]+0 307798 Sin[Ton]+0 0513136 Sin[WF]-0 0216832 Sin[W]				

B. Stability and Optimization Process of SR

Based on the stability graphs of the surface roughness parameter (Fig. 3), we can say that model is stable up to +10%. After +10%, it is obvious that a dramatic deviation occurs. Stability in the negative direction occurs up to -1%. After the stability calculations are done, optimum values of surface roughness are determined by using a derivative-free optimization algorithm, and results are presented in Table 4.

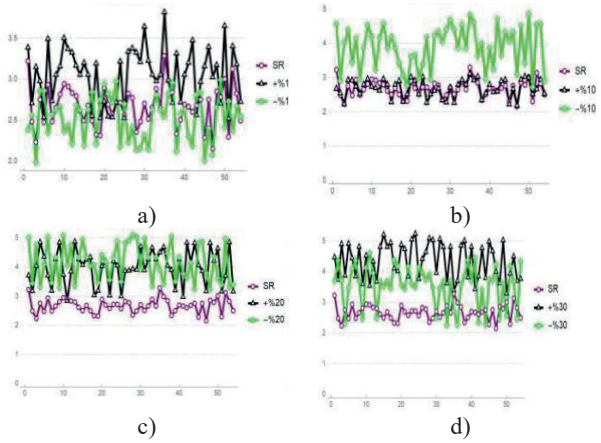


Figure 3: Stability graphs of the multiple nonlinear regression model for system output SR: a) $\pm 1\%$, b) $\pm 10\%$, c) $\pm 20\%$, d) $\pm 30\%$.

Table 4: Optimization results of the Simulated Annealing (SA) method.

Method	Optim. SR	T_{on} (µs)	T_{off} (µs)	IP (Ampere)	SV (Volt)	WF (m/min)	WT (gr)
SA	2.06	112	56	120	60	4	950

Results showed that it has been a significant increment of surface roughness (approximately 22%) with the rising of pulse on time from 112 to 120. Another important parameter affecting the surface roughness was peak current. As the peak current changed from 120 to 200 the surface roughness increased from 2.57 to 2.76. This situation depends entirely on the change in the intensity of the discharge energy. The higher peak current and pulse on-time result in a higher discharge energy level. This means longer machining time and stronger explosion [21]. This case is most probably result in double sparking and localized sparking [22]. The surface of the material to be rougher can be caused by both double sparks and re-sparks. Here the first phase sparks have a much more important effect in removing material from the surface, while the subsequent sparks have a secondary effect on the roughness as they act more on debris and ruptured particles[22, 23]. These sparks cause craters on the material surface. The roughness is directly proportional to the size of these craters. The higher the discharge energy, the stronger sparks, and the deeper craters are formed, which affects the surface roughness negatively. With less discharging energy, smoother surfaces occur because the diameter of the craters expands and their depth decreases. When we examine it in terms of its effect on roughness, the increase in spark gap

voltage shows a positive effect. The spark gap setting voltage is directly related to the discharge time. Increasing the spark gap voltage lengthens the discharge time. The longer the discharge time, the slower the machining speed, and the more stable the processing. This provides better surface roughness [23, 24]. Surface Roughness occurs as a result of erosion craters during machining. It is important to control discharging energy in order to obtain a flatter crater. This can only be achieved by choosing the optimal cutting parameters such as pulse on time, pulse off time, peak current, spark gap voltage, etc. that affect discharge energy. The results demonstrated the correct applicability of the stochastic optimization method in determining the most suitable cutting parameters in the WEDM process. As a result of the application of the Simulated Annealing Method, the process parameters providing the best surface roughness were determined. Surface roughness was calculated 2.06µm at pulse on time (112µs), pulse off time (56µs), peak current (120A), wire feed (4m/min), spark gap voltage (60volt), and wire tension (950gr).

IV. CONCLUSION

In this study, multiple nonlinear regression models that are suitable for phenomena were created. Calculated R^2, $R^2_{adjusted}$, $R^2_{training}$, and $R^2_{testing}$ values for the FOTN regression model were meaningful. Also, the stability of SR models is investigated to make sure that the model is not prone to deviate. Simulated Annealing method has been applied systematically to optimize WEDM's parameters. The optimal value of Surface roughness was found 2.06µm at pulse on time (112µs), pulse off time (56µs), peak current (120A), wire feed (4m/min), spark gap voltage (60volt), and wire tension (950gr) which are the best fitness values of cutting parameters of titanium plate by WEDM. This study has been a preliminary study to optimize other important output parameters such as Machining rate, Dimensional Deviations, Wire Wear Ratio, etc. in WEDM operation and is also important in terms of showing the applicability of Stochastic Optimization Methods.

REFERENCES

[1] A Kumar, V Kumar, J Kumar, "Multi-response optimization of process parameters based on response surface methodology for pure titanium using WEDM process", *The International Journal of Advanced Manufacturing Technology*, vol 68, pp 2645–2668, 2013

[2] H Adel, G E Hofy, "Advanced Machining Process", pp 1-10

[3] E O Ezugwu, Z M Wang, "Titanium alloys and their machinability," A Review *J Mater Process Technol,* vol 68, pp 262–274, 1997

[4] Y C , Lin, B H Yan, Y S Chang, "Machining characteristics of titanium alloy (Ti–6Al–4V) using a combination process of EDM with USM, " *J Mater Process Technology*, vol 104, pp 171– 177, 2000

[5] R Mukherjee, S Chakraborty, S Samanta, "Selection of wire electrical discharge machining process parameters using non-traditional optimization algorithms", Applied Soft Computing, vol 12, pp 2506-2516, 2012

[6] J T Black, R A Koher, *Materials and processes in manufacturing, 10th Ed.*, Hoboken, NJ: John Wiley and Sons, 2008

[7] P Shandilya, P K Jain, N K Jain, " Parametric optimization during wire electrical discharge machining using response surface methodology", *Procedia Engineering*, vol 38, pp 2371- 2377, 2012

[8] D Scott, S Boyina, K P Rajurkar, "Analysis and optimization of parameter combination in wire electrical discharge machining", *Int. J. Prod. Res.*, vol 29, pp 2189–2207, 1991

[9] T A Spedding, Z Q Wang, "Parametric optimization and surface characterization of wire Electrical discharge machining process", *Precis. Eng.* Vol 20, pp 5–15, 1997

[10] N Tosun, H Pihtili, "The effect of cutting parameters on wire crater sizes in wire EDM", *The International Journal of Advanced Manufacturing Technology*, vol 21, pp 857–865, 2003

[11] F Han, J Jiang, Y Dingwen, "Influence of machining parameters on surface roughness in finish cut of WEDM", *International Journal of Advanced Manufacturing Technology*, vol 34, pp 538-546, 2007

[12] S Singh, H K Kansal, P Kumar, "Parametric optimization of powder mixed electrical discharge machining by response surface methodology", *Journal of Materials Processing Technology,* vol 169, pp 427–436, 2005

[13] Z Kocabaş, "An Application and interpretation of second order response surface model", *Ankara Üniversitesi Tarım Bilimleri Dergisi,* vol 7 pp 121- 128, 2001

[14] M A Tekındal, H Bayrak, B Ozkaya, Y Genc, "Box- Behnken experımental design in factorıal experiments: The Importance Of Bread For Nutrıtıon", *Turkish Journal of Field Crops*, vol 17, pp 115-123, 2012

[15] M Sreenivasa Rao N Venkaiah, "A modified cuckoo search algorithm to optimize Wire-EDM process while machining Inconel-690J", *Braz. Soc. Mech. Sci. Eng*, vol 39, pp 1647–1661, 2017

[16] https://medium com/greyatom/logistic-regression-89e496433063

[17] http://www iasri res in/ebook/EB_SMAR/ebook_pdf%20files/Manual%20IV/1- Nonlinear%20Regression pdf

[18] D Pham, D Karaboga, *Intelligent optimization techniques, Genetic Algorithms, Tabu Search, Simulated Annealing And Neural Networks*, London: Springer Science and Business Media, 2012

[19] S Kirkpatrick, C D Gelatt, M P Vecchi, "Optimization by Simulated Annealing Science", New Series, vol 220, pp 671-680, 1983

[20] W D Callister, D G Rethwisch, *Materials Science and Engineering 9e : SI Version,* John Wiley & Sons Inc, 2014

[21] R T Yang, C J Tzeng, Y K Yang, M H Hsieh, "Optimization of wire electrical discharge machining process parameters for cutting tungsten", *Int J Adv Manuf Technol*, vol 30, pp 135–147, 2012

[22] R Ramakrishnan, L Karunamoorthy, "Surface roughness model for CNC wire electro-discharge machining", *Manuf Technol Toda,* pp 8–11, 2004

[23] S Kuriakose, M S Shunmugam, "Multi-objective optimization of the wire-electro discharge machining process by Non-Dominated Sorting Genetic Algorithm", *Journal of Materials Processing Technology,* vol 170, pp 133–14, 2005

[24] B J Ranganath, C P S Prakash, "Wire failure analysis in the WEDM process", *Manuf Technol Toda*, pp 19–24, 2005

Dynamic Simulation of Parabolic Trough Collector Integrated with Two-stage Organic Rankine Cycle by EES and Trnsys Software

Fathollah Pourfayaz[*,1] and Peyman Pourmoghadam[2]

[1] Department of Renewable Energies and Environment, Faculty of New Sciences and Technologies, University of Tehran, Tehran/Iran, pourfayaz@ut.ac.ir (Corresponding Author)
[2] Department of Renewable Energies and Environment, Faculty of New Sciences and Technologies, University of Tehran, Tehran/Iran, peyman.pourmoghadam@gmail.com

Abstract **- The organic Rankine cycles (ORC) are known as the reliable and low-temperature cycles. Also, solar energy is broadly considered one of the most exciting renewable energy for the future. So, the combination of the ORC and solar energy for producing electricity can lead to reducing fossil fuel consumption and diminish CO_2 emissions. Since the intensity of the solar radiation is different during a day, for precisely examining the performance of the system, the simulation should be done dynamically. In this regard, in this paper, the transient performance of the parabolic trough collector (PTC) cycle, which is combined with the two-stage ORC, is assessed. The main components of the system are PTC, a thermal storage system, ORC, auxiliary heater, and a controller, which shuts off the pump of the solar cycle when the radiation is low. Engineering Equation Solver (EES) and Trnsys software have been used to model the proposed system. The desired unit has been simulated during one year for a building which is located in Tehran. The results show that the energy efficiency of the whole cycle, which is used R245fa and R134a as the organic fluid, is 13.18%. Also, the highest energy efficiency and the solar fraction of the cycle occur in Jun with 13.28% and 73.46%, respectively, for the base case. The effect of the collector area, the storage tank volume, and the setpoint temperature of the auxiliary heater has been investigated on the performance of the system.**

Keywords **– ORC, PTC, EES, Trnsys, Dynamic.**

I. Introduction

Due to the increase in CO_2 emission and global warming, as well as the limited sources of fossil fuels, the use of renewable energy is needed more than ever. Solar energy seems to be more reliable and available than other renewable energies with respect to weather and location conditions. Due to the vast technology used for solar systems, parabolic trough collectors (PTCs) are more popular than other collectors owing to their use and performance on small and large scales [1]. On the other hand, organic Rankine cycles (ORCs) are considered as one of the most prevalent and competitive systems, owing to their acceptable operation temperature, to convert solar energy into electricity [2]. In this regard, Chai and Jung [3] investigated the performance of the 1 kW solar ORC system for a village in Malaysia. The R245fa and therminol VP1 were used as the organic and solar cycle fluid, respectively. Also, Heng et al. [4] conducted the performance of the 50 kW solar ORC, which used PTC with the 1154 m^2 aperture area. The efficiency of the proposed system was reported as around 8%. Yang et al. [5] examined the effect of the solar unit's stable output under variable radiation for one day. By keeping constant the temperature and flow rate of the heat transfer fluid, the unit's maximum efficiency reached 17.9%, which was 4.2% higher than the unstable output operating mode. Zhao et al. [6] investigated the effect of cloudy sky conditions on the performance of the cycle. It was illustrated that a short period (5 min) cloud blockage of the sun had no significant impact on the system's performance. Li et al. [7] evaluated a solar cascade cycle, which used steam and organic Rankine cycle as the top and bottom of the cycle, respectively. In this system, the screw expander was used to produce power, which can operate at lower pressure and temperature rather than steam turbines. The energy efficiency of this system was reported as about 15.67%. Kerme and Orfi [8] investigated the effect of turbine inlet temperature on the performance of the system. It was found that raising the turbine inlet temperature led to increasing the unit's electrical efficiency and output power. Freeman et al. [9] compared a PTC and evacuated tube collector (ETC) for the UK's residential house. It was found that the average electrical output power of the PTC, which was used tracker system, was very similar to the ETC. One of the important parameters of solar collectors is their array configuration. Zhang et al. [10] showed that the parallel array produced more electrical power in comparison with series one. Given the fact that solar systems can not work all day long, using the storage tank can increase their usage time. In this regard, Jafari mosleh and Ahmadi investigated the effect of phase change material (PCM) storage unit on the cycle's performance. Results showed that the solar fraction enhanced 90.5% compared to the system without the PCM storage unit. Caldino-Herrera et al. [11] investigated the thermal storage tank's effect on the small-scale PTC-ORC cycle, which used R245fa as the ORC working fluid. They claimed that using the storage tank can store up thermal load up to 11 hours. The proposed system had the 8.02% energy efficiency. Tzivanidis et al. [12] used a 300 m^3 storage tank for

a 25000 m^2 PTC solar field, which produced 1 MW power. The results showed that the internal return rate and the payback period of the desired system was 13.46% and 9 years, respectively. According to the importance of the ORC's working fluid on the performance of the system, Pourmoghadam and Pourfayaz [13] compared and optimized several working fluids for the cascade ORC. It was shown that R134a had the best thermal efficiency.

In this study, the transient simulation of solar ORC, which is integrated with PTC, is conducted for a building in Tehran. The desired system is comprised of PTC, ORC, storage tank, auxiliary heater, and the controller. Although the use of the auxiliary heater and the storage tank diminishes the temperature's fluctuation, the value of power generation as well as the operation of the other sections of the system, which is calculated over a year, can be very useful for designing the real configuration.

II. System Description

The simplified block diagram of the proposed system is shown in Figure 1. The operation of the cycle is as follows. The heat generated by the PTC is stored inside the storage tank. The DOWTHERM Q heat transfer fluid is used as the thermal oil of the solar cycle [14]. It should be noted that the solar cycle pump is equipped with the controller, which turns off the pump in case of insufficient or absence of radiation. After that, the outlet fluid from the storage goes to the auxiliary heater. If its temperature is less than the required value, it is heated up to the needed temperature. It then enters the heat exchanger and converts the high-temperature ORC's fluid to the superheated stream, which uses R245fa as the working fluid. In the ORC, the superheated stream rotates turbine blades to move the power shaft and generates electricity. After producing power, the turbine outlet stream goes to the condenser, which is also the low-temperature cycle's heat source. R134a is used as the operating fluid of the low-temperature cycle. Like the high-temperature cycle, the superheated steam causes the turbine blade to rotate, leading to producing electricity.

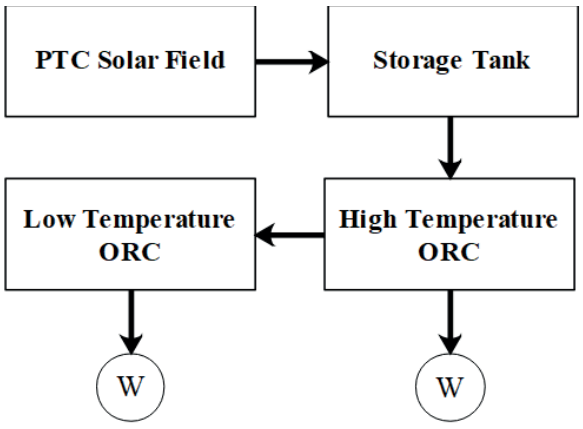


Figure 1: The schematic block diagram of the solar two-stage ORC

III. Simulation and Assumptions

A. Dynamic Simulation Tools

In this study, due to the transient simulation, Trnsys software is used to investigate the desired system's behavior during a year. Simulation is performed according to Tehran's weather data and solar radiation, which are illustrated in Figure 2. The type TMY2 is used to read the weather data of Tehran. Also, the time step of the simulation is chosen to be an hour. Moreover, Engineering Equation Solver (EES) is employed to model the ORC components, which are not available in the Trnsys.

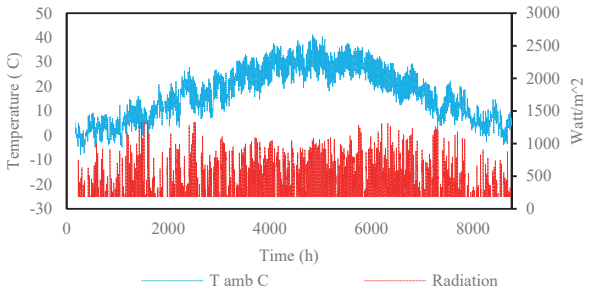


Figure 2: Annual radiation and the temperature of Tehran

B. Solar Field

According to the building's electrical demands and solar radiation condition in Tehran, 4 PTCs in two rows are considered. Each PTC has a 70 m^2 aperture area with 23 m long and 3 m span. The simulated solar collector system is equipped with a single-axis tracking to trace sunlight during day time. The properties of the PTC are listed in Table 1.

Table 1: Specification of parabolic trough collector

Parameter	Value
Length (m)	20
Width (m)	3.5
Number of Collector in Series	2
Number of Collector in Parallel	2
Total Aperture Area (m^2)	280
Concentration Rato	35
Intercept Factor	0.93
Collector Flow Rate (kg/s)	3.33
Fluid Specific Heat (kj/kg.K)	2.5

C. Organic Rankine Cycles

The ORC comprises turbines, pumps, heat exchangers, and a condenser, which are modeled in the EES and then integrated into the Trnsys software environment. Isentropic efficiency was considered for turbines and pumps. Also, the condenser outlet fluid is assumed to be a saturated liquid. The properties and assumptions of the ORC modeling are listed in Table 2.

Table 2: Specifications of ORC

Parameter	Value
Heater Setpoint Temperature (C)	155
High-Temperature ORC fluid	R245fa
High-Temperature ORC Flow Rate (kg/s)	0.55
High-Temperature ORC Pressure (Kpa)	2665
Low-Temperature ORC fluid	R134a
Low-Temperature ORC Flow Rate (kg/s)	0.55
Low-Temperature ORC Pressure (Kpa)	2200
Turbine Isentropic Efficiecny	0.85
Turbine Mechanical Efficiency	0.85
Pump Isentropic Efficiecny	0.75
Pump Mechanical Efficiency	0.75
Outlet Quality of the Condenser	0

IV. Performance Criteria and Validation

A. Turbine

In the case of modeling turbines, which have been done in EES, turbines are considered adiabatic, and the turbine generator's loss is assumed to be 0.85. The isentropic efficiency and turbine outlet power are defined in equations 1 and 2, respectively.

$$\eta = \frac{h_{in,Turbne} - h_{out,Turbne}}{h_{in,Turbne} - h_{out,is,Turbne}} \quad (1)$$

$$W_{Turbne} = \dot{m}_{Turbine} \times (h_{in,Turbne} - h_{out,Turbne}) \times 0.85 \quad (2)$$

B. Pump

In the modeling of Pumps, which have been done in EES, Pumps are considered adiabatic. The isentropic efficiency and the pump's consumption power are defined in equations 3 and 4, respectively.

$$\eta = \frac{h_{out,is,Pump} - h_{in,Pump}}{h_{out,Pump} - h_{in,Pump}} \quad (3)$$

$$W_{Pump} = \frac{\dot{m}_{Pump} \times (h_{out,Pump} - h_{in,Pump})}{0.75} \quad (4)$$

C. Solar Fraction

The solar fraction is usually defined to evaluate the performance of the collector. The solar fraction is the ratio of the amount of energy supplied by the sun to the required cycle's total energy. In this study, the solar fraction is defined as:

$$SF = \frac{Q_{Solar}}{Q_{Solar} + Q_{Auxilary\ boiler}} \quad (5)$$

The total efficiency of the cycle is defined as:

$$\eta_{Cycle} = \frac{\sum W_{Turbine} - \sum W_{Pump}}{Q_{Solar} + Q_{Auxiliary\ boiler}} \quad (6)$$

D. Validation

Before modeling the proposed system, the two-stage ORC modeled in EES was compared with the Kosmadakis et al. [15] results. Table 3 shows a good agreement between simulation results and the operating data in Ref [15]. The reason for these tiny errors is the small difference between the proposed and Ref [15] cycles.

Table 3: Comparision the ORC developed model in EES with the results of Ref [15]

	Data in Ref [15]	Simulation Data	Error (%)
High-Temperature ORC power production (kW)	8.5	8.91	4.82
Low-Temperature ORC power production (kW)	6.9	6.48	6.08
High-Temperature ORC condenser heat rejection (kW)	103	92.60	10.09
Low-Temperature ORC condenser temperature (C)	35	35	0

V. Results

The total radiation on the PTC's tilted surface and the horizontal's total radiation are illustrated in Figure 3 for a random day in Jun. As can be seen, the radiation increases from zero in the morning to its maximum value at noon and then it reduces again to reach zero at sunset. The sun's radiation patterns on all days are the same as mentioned, and only their amounts are different. Furthermore, the difference between the radiations indicates the amount of increase in radiation, which was received by the PTC due to the single-axis tracking system. How the control system works is also shown in Figure 3. At sunrise, as the solar intensity increases, it turns on the pump, and at sunset, it shuts off the pump as the radiation diminishes.

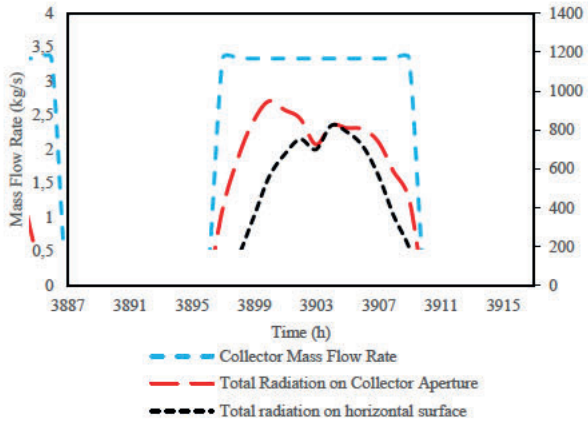


Figure 3: Collector mass flow rate and radiation on collector aperture and horizontal surface in Jun

As can be seen in Figure 4, variation in auxiliary heater consumption and the useful energy received by the PTC have been shown. It is observed that as the radiation is high, the auxiliary power is reduced. Even during part of the day, it was turned off. Also, in some parts of the day, neither the PTC nor the auxiliary heater work. At this time, the energy stored in the storage tank provides the energy needed for the cycle.

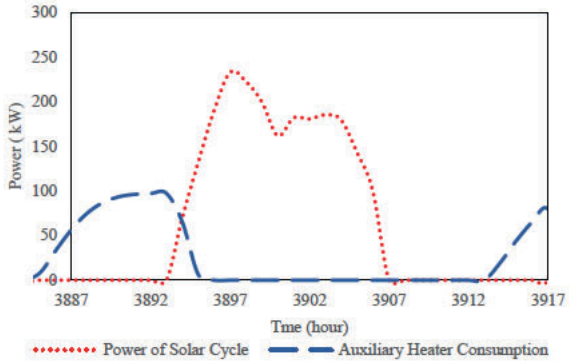


Figure 4: The auxiliary heater consumption and useful energy received by PTC

The monthly solar fraction and efficiency of the base case and the monthly turbine power production, heater consumption, and the PTC useful energy gain are shown in Figure 5 and Figure 6, respectively. As can be seen, the closer we got to the summer season, the more radiation the PTC received. Consequently, the heater consumed less work, which led to increasing solar fraction. According to the constant monthly value of turbine power production, the monthly total energy efficiency did not change significantly. The maximum solar fraction occurred in Jun, which is about 73.46%. Furthermore, the highest amount of energy received by PTC happened in July with 54836.34 kWh.

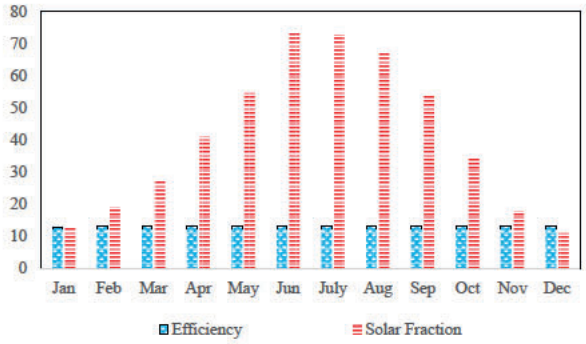


Figure 5: The monthly Efficiency and solar fraction for the base case condition

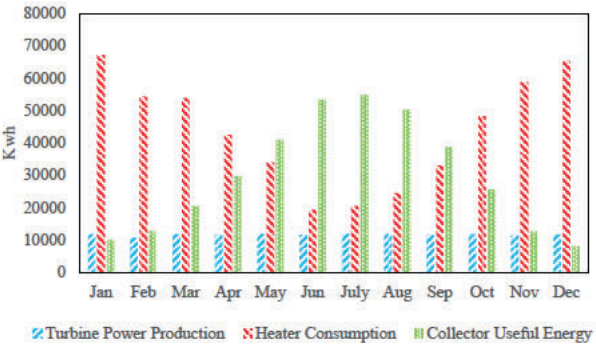


Figure 6: The monthly power production, heater consumption, collector useful energy for the base case condition

The effect of increasing the aperture area of the PTC on the solar fraction, efficiency, PTC useful energy gain, heater consumption, and power production is illustrated in Table 4. By increasing the area from 280 m^2 to 380 m^2, the total solar fraction was enhanced from 40.67% to 48.38%. Since the amount of power production remained constant and the value of the input energy to the cycle was increased, which is the summation of solar energy and heater consumption, efficiency reduced from 13.18% to 11.6%.

The effect of increasing the auxiliary heater's setpoint temperature on the solar fraction, efficiency, PTC useful energy gain, heater consumption, and power production is mentioned in Table 4. By increasing the setpoint temperature from 155 C to 175 C with the 280 m^2 aperture area, which is the base case in this study, the total heater consumption increased from 521369 kWh to 842237 kWh. Consequently, the solar fraction diminished from 40.67% to 29.68%. Although in case 3, the amount of power production was increased, its value was not enough to overcome the increment of heater consumption. So the efficiency decreased from 13.18% to 10.01%.

Table 4: The total results of the sensitivity analysis on the PTC-ORC

	Case 1	Case 2	Case 3
Aperture area (m^2)	280	380	280
Heater setpoint temperature (C)	155	155	175
PTC useful energy gain (kWh)	357429	485109	357429
Power production (kWh)	139665	140030	143779
Heater consumption (kWh)	521369	517515	842237
Solar fraction	40.67	48.38	29.68
Efficiency	13.18	11.6	10.01

The effect of increasing the tank volume from 50 m^3 to 75 m^3 was investigated on the solar fraction, efficiency, and power production. Since its changes approximately did not affect the system's performance, no result was presented.

VI. Conclusion

In this study, the transient simulation of the two-stage ORC, which was coupled with the PTC, is investigated. Some of its essential results are as follows.

- Although by increasing the PTC area, the energy received by PTC increases and the power consumption of heater decreases, the efficiency of the whole cycle reduces. However, solar fraction increases dramatically.
- The use of the auxiliary heater leads to decreasing the effect of the radiation's fluctuation on electricity production by keeping the system's energy level constant. It should be noted that, by using the proper energy storage tank, the heater's consumption can be diminished.

References

[1] H Hoseinzadeh, A Kasaeian, and M B Shafii, "Exergoeconomic optimization and sensitivity analysis of a commercial parabolic trough collector for the climate of Tehran, Iran," *Energy Science & Engineering,* vol 7, no 6, pp 2950-2965, 2019, doi: 10 1002/ese3 472

[2] O Aboelwafa, S -E K Fateen, A Soliman, and I M Ismail, "A review on solar Rankine cycles: Working fluids, applications, and cycle modifications," *Renewable and Sustainable Energy Reviews,* vol 82, pp 868-885, 2018/02/01/ 2018, doi: https://doi org/10 1016/j rser 2017 09 097

[3] C Chai and H -c Jung, "Sizing a parabolic trough collector for a micro solar organic rankine cycle," *Jurnal Teknologi,* vol 81, 01/22 2019, doi: 10 11113/jt v81 12304

[4] S Y Heng, Y Asako, T Suwa, L K Tan, N B Sharifmuddin, and J O Kamadinata, "Performance of a small-scale solar cogeneration system in the equatorial zone of Malaysia," *Energy Conversion and Management,* Article vol 184, pp 127-138, 2019, doi: 10 1016/j enconman 2019 01 059

[5] J Yang, J Li, Z Yang, and Y Duan, "Thermodynamic analysis and optimization of a solar organic Rankine cycle operating with stable output," *Energy Conversion and Management,* Article vol 187, pp 459-471, 2019, doi: 10 1016/j enconman 2019 03 021

[6] J Ni *et al.*, "Dynamic performance investigation of organic Rankine cycle driven by solar energy under cloudy condition," *Energy,* Article vol 147, pp 122-141, 2018, doi: 10 1016/j energy 2018 01 032

[7] J Li, P Li, G Pei, J Z Alvi, and J Ji, "Analysis of a novel solar electricity generation system using cascade Rankine cycle and steam screw expander," *Applied Energy,* Article vol 165, pp 627-638, 2016, doi: 10 1016/j apenergy 2015 12 087

[8] E D Kerme and J Orfi, "Exergy-based thermodynamic analysis of solar driven organic Rankine cycle," *Journal of Thermal Engineering,* Article vol 1, no 5, pp 192-202, 2015, doi: 10 18186/jte 25809

[9] J Freeman, K Hellgardt, and C N Markides, "An assessment of solar-powered organic Rankine cycle systems for combined heating and power in UK domestic applications," *Applied Energy,* Article vol 138, pp 605-620, 2015, doi: 10 1016/j apenergy 2014 10 035

[10] H Zhang, B Lei, T Yu, and Z Zhao, "Exergy analysis of two kinds of solar-driven cogeneration systems in Lhasa, Tibet, China," *International Journal of Photoenergy,* Article vol 2018, 2018, Art no 6702049, doi: 10 1155/2018/6702049

[11] U Caldiño-Herrera, L Castro, O A Jaramillo, J C Garcia, G Urquiza, and F Flores, "Small Organic Rankine Cycle Coupled to Parabolic Trough Solar Concentrator," in *Energy Procedia*, 2017, vol 129, pp 700-707, doi: 10 1016/j egypro 2017 09 097

[12] C Tzivanidis, E Bellos, and K A Antonopoulos, "Energetic and financial investigation of a stand-alone solar-thermal Organic Rankine Cycle power plant," *Energy Conversion and Management,* Article vol 126, pp 421-433, 2016, doi: 10 1016/j enconman 2016 08 033

[13] P Pourmoghadam and F Pourfayaz, "Optimization and analysis of a power generation cascade cycle with the Genetic algorithm," presented at the The 28th Annual International Conference of the Iranian Association of Mechanical Engineers, 2020 [Online] Available: https://civilica com/doc/1029467

[14] *D. C. Company, DOWTHERM Q Heat Transfer Fluid, Product Technical Data, ed.USA, 1997*

[15] G Kosmadakis, D Manolakos, S Kyritsis, and G Papadakis, "Simulation of an autonomous, two-stage solar organic Rankine cycle system for reverse osmosis desalination," *Desalination and Water Treatment,* vol 1, pp 114-127, 08/03 2012, doi: 10 5004/dwt 2009 115

Effects of Nozzle Throat Insert Geometry on Stresses

M.SAHAN[1] and E. ACAR[2]

[1] TOBB University of Economics and Technology, Ankara/Turkey, m.sahan@etu.edu.tr
[2] TOBB University of Economics and Technology, Ankara/Turkey, acar@etu.edu.tr

Abstract - **One of the most important issues in solid fuel rocket motor design is that the designed nozzle can function under thermomechanical effects. Throat insert cracking often occurs due to thermal stresses and can cause local material loss around the crack. Any shape change on the inner surface of the nozzle directly affects motor performance negatively. Numerous studies have been conducted on the thermomechanical response of solid fuel rocket motor nozzles. In these studies, many factors related to the geometrical properties and material properties of rocket motor nozzles were taken into consideration. In this paper, the boundary conditions required for thermomechanical analysis were obtained by using a one-dimensional nozzle flow solver code. The obtained thermal and structural boundary conditions were applied to the structural analysis model created using ANSYS software, and the effects of various geometrical properties of the nozzle throat insert on the stresses on the nozzle throat insert were investigated. It is found that allowing radial gap between the insert and the insulation material is important to reduce compression stresses. Additionally, 10mm more thinner insert design showed an improvement in the critical minimum principle stress value for this geometry. However, when the insert thickness value is decreased further, this time a sudden increase in the maximum principle stress value is observed. This paper provides insights into the effects of nozzle throat insert geometry in solid propellant rocket motor nozzles on stresses and provides an approach to obtain more reliable design.**

Keywords - **Solid Propellant Rocket Motors, Nozzle Throat Insert Design, Thermomechanical Analysis**

I INTRODUCTION

A. Solid Rocket Motor

SOLID propellant rocket motor is a type of rocket motor whose fuel and oxidizer are in solid phase, and these solid propellant rocket motor systems are used as accelerators on missiles, satellite launch systems and model rockets. The main function of solid rocket motors is to reach the ammunition to the desired target.

Rocket motors principally consist of four main parts: (i) case, (ii) igniter, (iii) propellant, (iv) nozzle. Motor case contains solid propellant and resistant to internal pressure while the rocket operates. The igniter is generally a pyrotechnic device which starts the combustion when an electrical signal is received. Solid propellant, usually attached to the inner wall of the motor case, provides necessary energy for propulsion and takes a large part of case volume before ignition. After ignition occurs, combustion starts on the solid propellant surface, then the solid propellant turns into hot combustion products. As a result of combustion, large amounts of exhaust gas are produced at high temperature and pressure. Nozzle is used as a device in which the chemical energy of the propellant is converted into the kinetic energy of the gas at high temperature and pressure. Exhaust gas at high temperature and pressure is passed through the nozzle that accelerates the flow. Then, thrust is generated by the impulse force according to Newton's third law of motion.

B. Solid Rocket Motor Nozzle

Nozzle design is an iterative procedure, firstly the design process is completed according to the design parameters and constraints and then the design is checked with analysis, finally nozzle design validated by tests (see Fig. 1).

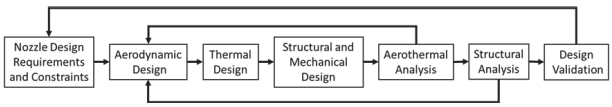


Figure 1. Basic Nozzle Design Process

Aerodynamic design phase is the geometrical design of entrance, throat and exit surfaces of nozzle that are exposed to exhaust gases and directly affects motor performance. The main purpose in thermal design is to maintain the nozzle aerodynamic design as much as possible and to protect the structural parts from high temperatures. Thermal insulator and throat insert materials are used for this purpose. The aim of structural and mechanical design is to determine the basic external structure which supports thermal materials and carries nozzle loads. In short, at the end of the design, the rocket motor nozzle basically consists of three layers: (i) ablation layer, (ii) thermal insulation layer, (iii) support layer.

C. Studies in the Literature

Numerous research and investigations are conducted to understand the thermomechanical response of rocket motor nozzle until today. Bartz (1957) suggests a simple equation for rapid estimation of rocket nozzle convective heat transfer coefficients. Kumar et al. (2005) thermal loads are more dominant in thermomechanical response of the nozzle than mechanical loads. Sun et al. (2016) examined the effect of structural gaps on the thermomechanical response of a typical nozzle using a finite element method. Sun et al. (2017) studied the thermo-structural response of the graphite throat using finite element analysis without considering chemical and mechanical ablation processes to determine the failure cause of a crack. Unlu et al. (2018) used Abaqus to analyze thermo-mechanical response of the nozzle of missile CASTOR II, propose design changes to reduce the maximum and minimum

principle stress levels. Nigar (2019) investigated several geometries for their crack propagation behavior.

In this paper, the effects of insert thickness/length radial structural gaps and material properties on stresses are investigated, with an aim to improve the design of the nozzle throat insert geometry.

II. Problem Definition

When the propellant is ignited, temperature and pressure begin to increase rapidly in the combustion chamber and nozzle. The surface temperature of the nozzle parts, which is exposed to flow, reaches 2500-3000°K. A large temperature gradient occurs inside the nozzle parts, because the external layer of the rocket nozzle is still at room temperature. In addition to stresses due to high pressure, this temperature gradient causes a thermal expansion that causes large thermal stresses. Finally, this phenomenon can cause cracks of nozzle throat insert and consequently failure of nozzle design. Any crack failure in the nozzle throat cause the solid rocket motor to fail its mission.

At the scope of this investigation the design variables of the nozzle throat insert are determined as follows: (1) Length of the nozzle throat inserts, (2) Radial thickness of the nozzle throat inserts, (3) Radial structural gap between insert and insulator.

III. Thermo-Mechanical Analysis

To perform thermomechanical analysis of a rocket motor nozzle, the following steps should be applied.

Creating the solid model of the nozzle design: Nozzle geometry including ablation layer, thermal insulation layer and support layer can be created by using mechanical design software.

Defining the temperature-dependent material properties: All properties of materials (e.g., density, thermal conductivity, thermal expansion coefficient, modulus of elasticity, Poisson's ratio, specific heat) should be temperature-dependent to obtain more accurate results because transient thermal analysis is handled.

Obtaining of 2-D axially symmetric finite element analysis model: The nozzle geometries are axially symmetrical. Therefore, 2-dimensional axially symmetric finite element model can be used to obtain accurate results in shorter time using a smaller number of mesh element and nodes.

Creating appropriate mesh and defining proper time step: Appropriate element and time step size should be determined to obtain results independent of time step and element size.

Identifying the contact status: Physical contacts between nozzle parts must be properly defined. Frictional and bonded contact status are used generally.

Determination of boundary conditions: Two types of boundary conditions should be determined to perform thermomechanical analysis of the nozzle. (A) Thermal Boundary Conditions: Free convection heat transfer between external nozzle surface and air and forced convection heat transfer between flow field and internal nozzle surface. (B) Structural Boundary Conditions: Distributed pressure on internal surface of nozzle because of flow field.

Running thermomechanical analysis and obtaining results: For the problem to be investigated, analyzes can be performed by ignoring the erosion and pyrolysis of the materials.

A. 1-D Nozzle Flow Simulation Code

This code is used for determination of nozzle flow properties and defining boundary conditions of thermomechanical analysis. To obtain boundary conditions for thermomechanical analysis, we use a 1-D flow simulation code to obtain flow properties along the nozzle axis. Flowchart of the 1-D nozzle flow simulation is shown in Fig. 2.

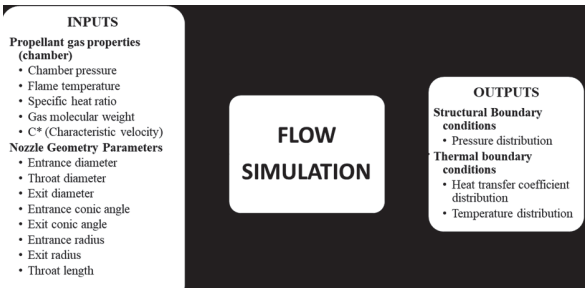


Figure 2. Flowchart of the 1-D Nozzle Flow Simulation

For 1-D isentropic flow, relationship between temperature, pressure, and Mach number in any section along the nozzle axis can be given as:

$$\frac{T_0}{T} = 1 + \frac{k-1}{2} Ma^2 \quad (1)$$

$$\frac{P_0}{P} = (1 + \frac{k-1}{2} Ma^2)^{\frac{k}{k-1}} \quad (2)$$

$$\quad (3)$$

where, is chamber temperature, P_0 is chamber pressure, A_t is nozzle throat area, A is the local area along the nozzle axis and k is the specific heat ratio of the gas.

By using of these equations, firstly, the area ratio of the relevant axial coordinate (A/A_t) and the local Mach number (Ma) are obtained. Using Ma, combustion chamber temperature and specific heat ratio, the local static temperature value (T) is obtained. The local static pressure value (P) is obtained from Ma, combustion chamber pressure and free heat rate. Then obtained pressure distribution is used as internal pressure structural boundary condition.

To calculate the forced convection heat transfer from the internal flow to the internal surface of the nozzle, the gas recovery temperature T_r must be provided as the temperature of the gas directly contacting the nozzle. The recovery temperature can be calculated from:

$$T_r = T\left[1 + Pr^{(1/3)} \frac{k-1}{2} Ma^2\right] \quad (4)$$

(5)

where, T and Ma are the gas static temperature and the Mach number in the local cross-sectional area.

For forced convection, the heat transfer coefficient can be calculated with the Bartz formulation:

$$h = \frac{C}{{d_t}^{0.2}} \frac{{c_p}{\mu}^{0.2}}{Pr^{0.6}} \left(\frac{\dot{m}}{A_t}\right)^{0.8} \left(\frac{d_t}{r_c}\right)^{0.1} \left(\frac{A_t}{A}\right)^{0.9} \sigma \quad (6)$$

[redacted] (7)

$$\mu = 11.83x10^{-8} x\, \bar{M}^{0.5} x\, T^{0.6} \quad (8)$$

where, h is the forced convection coefficient, C is a constant, D_t is throat diameter, c_p is constant volume specific heat of gas, μ is the viscosity coefficient of the gas, Pr is Prandtl number, ṁ is mass flow rate, r_c is radius of curvature, A_t is throat cross section area, A is local cross section area, T_w is the temperature of the wall, ' is temperature of the total temperature of the gas, k is the ratio of specific heat, Ma is local Mach number, σ is a modified parameter considering the vibration of the boundary layer, $\bar{M}$ is the gas molecular weight.

IV. Validation of the Finite Element Model

Thermomechanical analysis of the CASTOR IIA nozzle is first conducted to verify the thermomechanical analysis model by comparing the results of this study with the results available in the literature. Geometrical properties and design materials are shown Figures 3 and 4.

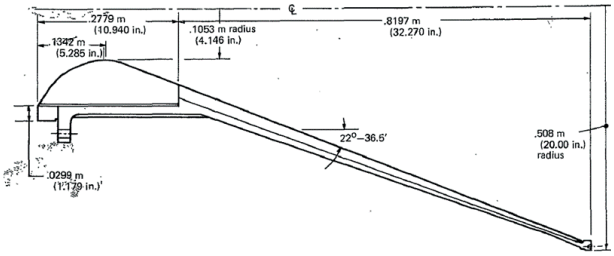


Figure 3: Geometrical properties of CASTOR IIA nozzle (courtesy of Shields, 1976)

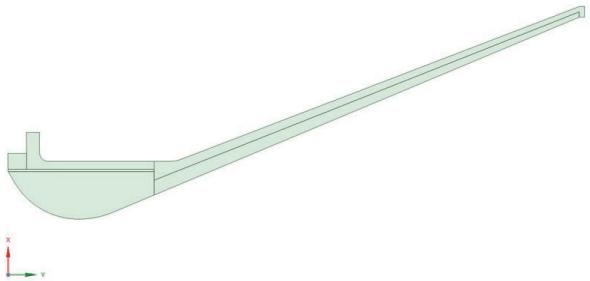
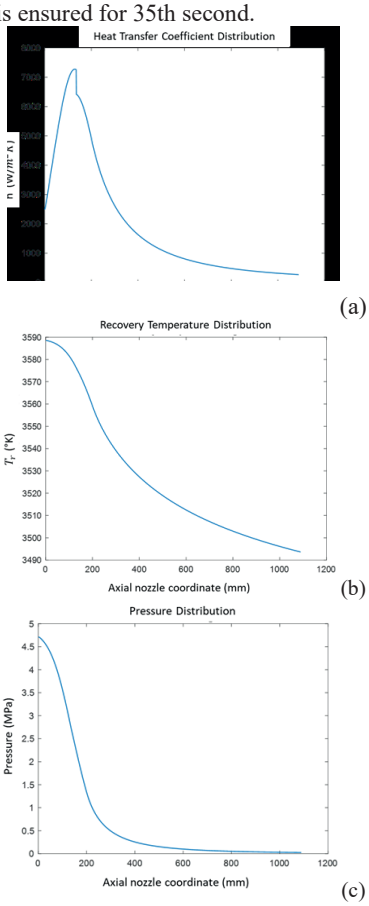


Figure 4: 2D model of CASTOR IIA nozzle

A. Comparison of The Pressure Distribution

By using flow simulation code, the properties of the flow along the inner contour of the nozzle are determined and used as the boundary conditions in the analysis model. The outputs of flow simulation is shown Fig 5. Additionally, pressure distribution which is obtained by using flow simulation code is shown red color on the pressure distribution graph which is provided from literature for CASTOR IIA nozzle. When the Figure 6 is examined, the same pressure distribution is ensured for 35th second.



(a) (b) (c)

Figure 5: Nozzle flow properties (a-c)

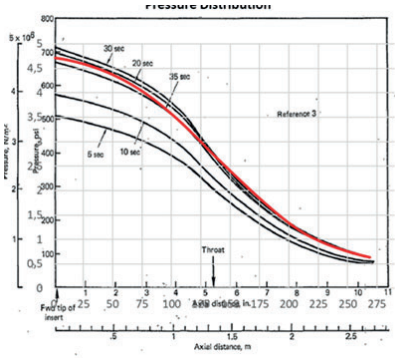
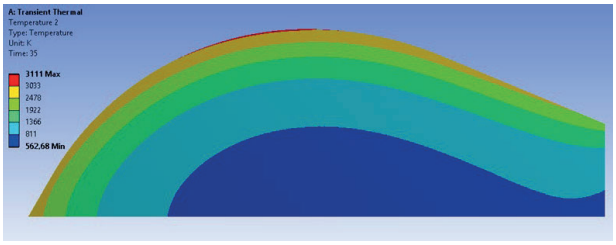


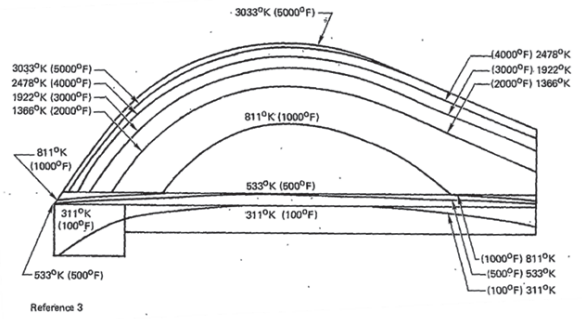
Figure 6: Pressure distribution and comparison for 35th sec. (red line is this study on literature result)

B. Comparison of The Temperature Distribution

Temperature distribution obtained in this study is found to be close to that of the literature (see Fig. 7). The maximum and minimum temperature values was around 3033 ºK and 562 ºK in literature, and they are computed around 3111 ºK and 533 ºK in this study.



(a) Temperature Distribution, 35th second

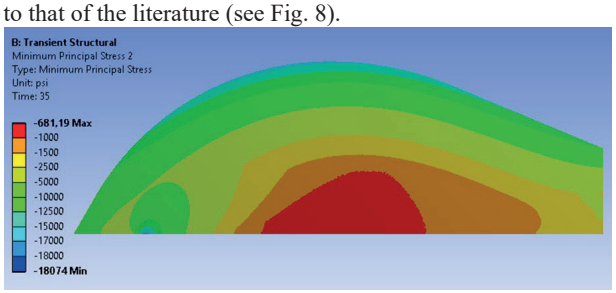


(b) Temperature Distribution, 35th second (Shields, 1976)

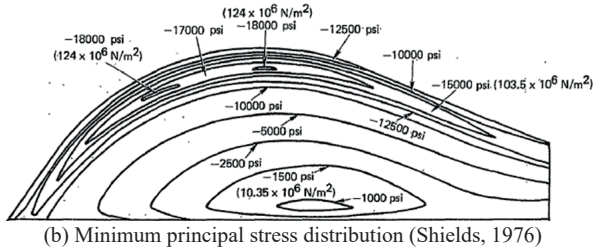
Figure 7: Comparison of temperature distribution (a: this study, b: literature)

C. *Comparison of the Minimum Principal Stress Distribution*

Minimum principal stress distribution is found to be close to that of the literature (see Fig. 8).



(a) Minimum principal stress distribution

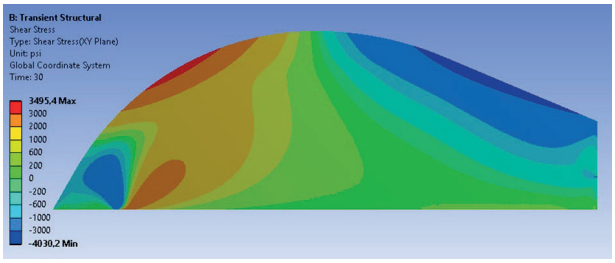


(b) Minimum principal stress distribution (Shields, 1976)

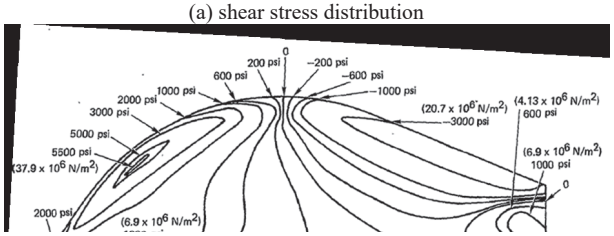
Figure 8: Comparison of minimum principal stress distribution at the 35th second (a: this study, b: literature)

D. *Comparison of the Shear Stress Distribution*

Shear stress distribution obtained in this study is also found to be close to that of the literature (see Fig. 9).



(a) shear stress distribution



(b) shear stress distribution (courtesy of Shields, 1976)

Figure 9: Comparison of shear stress distribution at the 30th second (a: this study, b: literature)

V. EFFECT OF VARIOUS PARAMETERS ON STRESSES

After completing the verification of the analysis method by using the Castor's model and results in the literature, our own nozzle design (see Fig. 10) is analyzed using the same method.

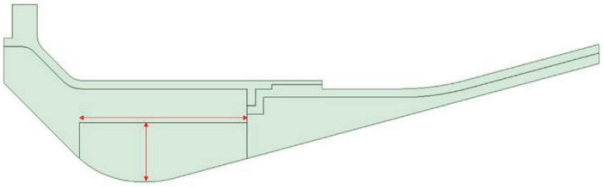


Figure 10. 2D model of examined nozzle

Nozzle dimensions and working conditions are given in Table 1. Applied boundary conditions and contacts are shown in Figure 11. Flow simulation results are shown in Figure 12.

Table 1. Nozzle dimensions and working conditions

	Parameter	Value
Nozzle Dimensions	Throat diameter	110mm
	Exit diameter	250mm
	Nozzle Length	357mm
Operating conditions	Pressure	7 5 MPa
	Temperature	3600 °K
	Operating time	5s

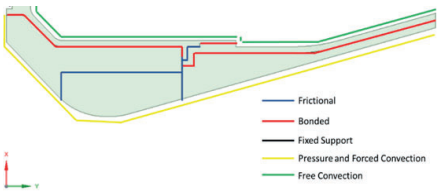


Figure 11. Boundary conditions and contacts

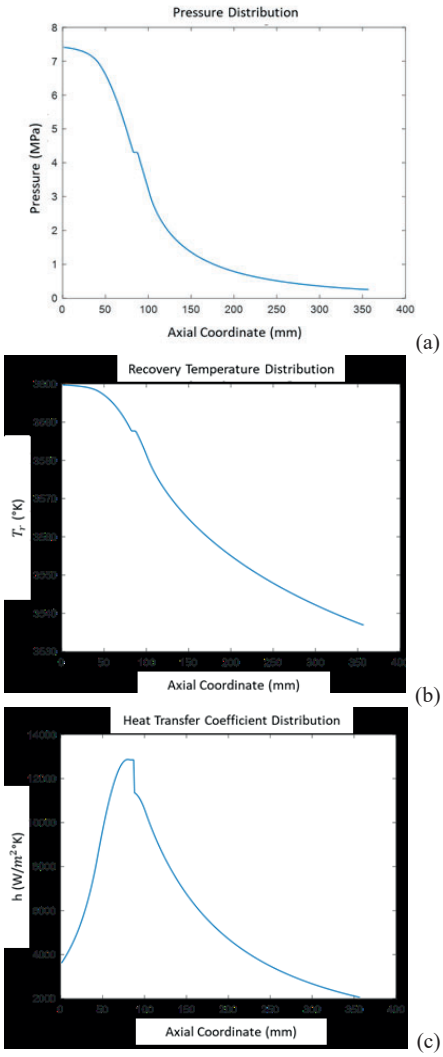


Figure 12: (a) Pressure distribution, (b) temperature distribution, and (c) heat transfer coefficient distribution

A. Mesh and Time Step Dependency

Four different element sizes (5mm, 2mm, 1mm, 0.5mm) and five different time steps (1s, 0.5s, 0.2s, 0.1s) is examined to obtain independent results. A path is created from the throat to the external surface of insert, then change in von Mises stress and temperature are investigated. We found that, when the element size is reduced to 1mm, the results are independent of the mesh and when the time step is reduced to 0.1 seconds, the results are independent of the time step. Due to page limitations, we skip mesh and time step dependency figures.

B. Reference Results

The minimum principle stress and the maximum principle stress results of the reference geometry are shown in Figures 13 and 14, and they will be used as a reference in comparisons. Considering that Graphite ATJ has a compressive strength of 80 MPa, the region shown on the front and behind contact surface of the insert can be considered critical in terms of compressive stress. Considering that Graphite ATJ has a tensile strength of 25 MPa, the region shown on the behind contact surface of the insert can be considered critical in terms of tensile stress.

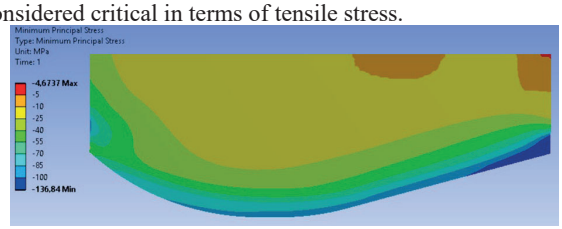


Figure 13. Minimum principle stress distribution on insert

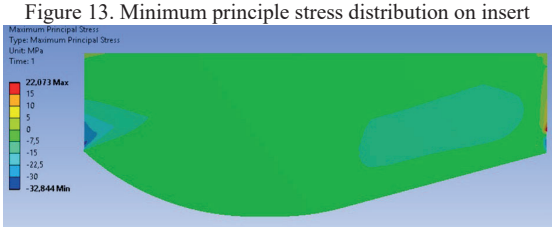


Figure 14. Maximum principle stress distribution on insert

C. Effect of Thicknesses of the Nozzle Throat Insert

When the insert is 10mm thinner (see Fig. 15), it is found that the critical minimum principal stress values are not affected much by this change, but the critical minimum principal stress value is decreased from 22.1 MPa to 19.3 MPa (see Figs. 16 and 17).

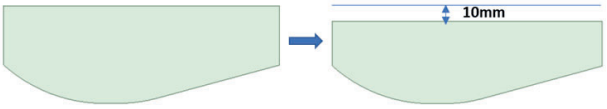


Figure 15. Applied geometrical change

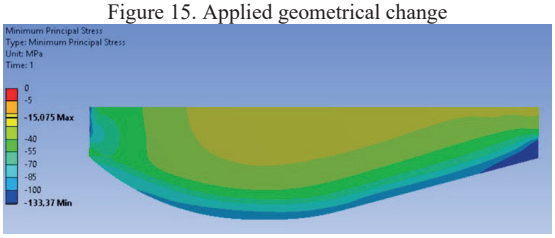


Figure 16. Minimum principle stress distribution on insert

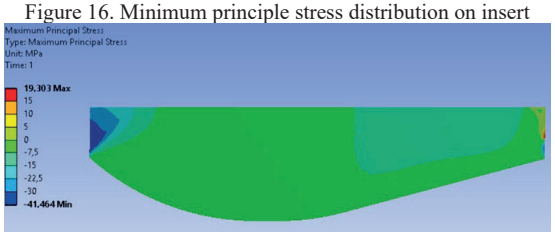


Figure 17. Maximum principle stress distribution on insert

When the insert is 15mm thinner, it is observed that the critical minimum principal stress value is decreased to 16 MPa, but a sudden drop of critical minimum principal stress values can be seen in the critical region at the front interface of insert (see Figs. 18 and 19). Geometries between 10 mm and 15 mm thinner geometries should be studied in more detail to determine at what thickness the minimum principle stress value decreases suddenly.

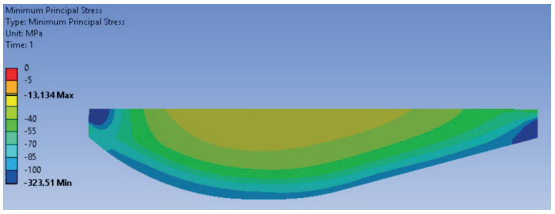


Figure 18. Minimum principle stress distribution on insert

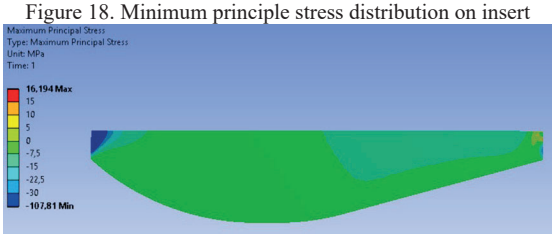


Figure 19. Maximum principle stress distribution on insert

D. *Effect of Length of the Nozzle Throat Insert*

The analyses are repeated with 10mm longer insert geometry (see Fig. 20). While the maximum and minimum principle stress values do not change much when the insert length is 10mm longer, it has been observed that the region under critical compressive stress on the front of the insert expands (see Figs. 21 and 22). It can be interpreted that when the insert length is shortened, the area where critical compression stresses are observed on the front surface of the insert will decrease.

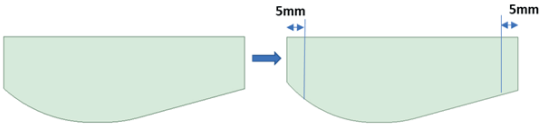


Figure 20. Applied geometrical change

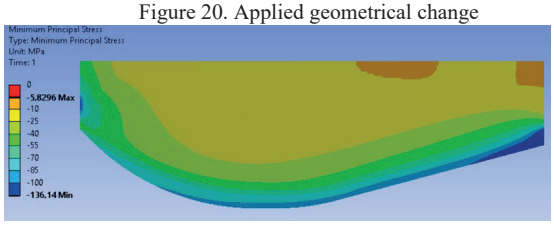


Figure 21. Minimum principle stress distribution on insert

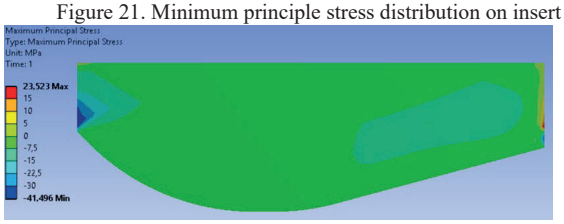


Figure 22. Maximum principle stress distribution on insert

E. *Effect of Radial Gap*

The analyzes are repeated with 0.25mm radial gap between insert and insulator material and the effect of radial gap on stresses was investigated. It is seen that compression stresses decrease but tensile stresses increase. If compression stresses are more critical than tensile stresses, allowing radial clearance can be used to improve the design (see Fig. 23 and 24).

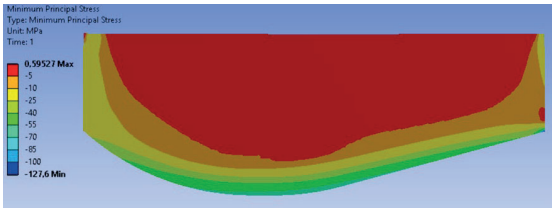


Figure 23. Minimum principle stress distribution on insert

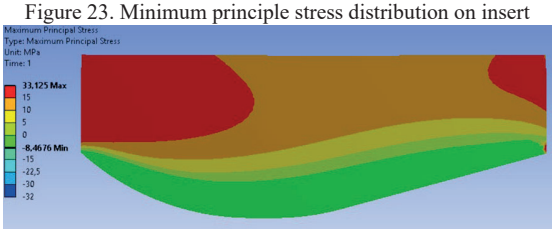


Figure 24. Maximum principle stress distribution on insert

VI. Concluding Remarks

To provide insights into obtaining more reliable solid propellant rocket motor nozzles design, this paper investigates the effects of nozzle throat insert geometry on stresses developed. The following conclusions were drawn:

The 10 mm thinner insert design showed an improvement in the critical minimum principle stress value. However, when the insert thickness value was decreased by 15 mm, this time a sudden increase in the maximum principle stress value was observed. In a future study, the thickness reduction between 10 and 15mm would be examined in more detail.

It was observed that increasing the insert length expands the critical region in terms of compressive stress at the front interface of the insert. In a future study, the effect of shortening the insert size would be explored.

Allowing radial space between the insert and the insulation material was found to be important in reducing compression stresses. However, it was observed that this radial gap affects tensile stress values negatively.

References

[1] D R Bartz A Simple Equation for Rapid Estimation of Rocket Nozzle Convective Heat Transfer Coefficients, Jet propulsion, 27 (1957) 49-51

[2] Kumar, R R ; Vinod, G ; Renjith, S ; Rajeev, G ; Jana, M K ; Harikrishnan, R Thermo-structural analysis of composite structures Mater Sci Eng A 2005, 412, 66–70

[3] L Sun, F T Bao, N Zhang, W H Hui, S Z Wang, N Zhang, H Deng Thermo-Structural Response Caused by Structure Gap and Gap Design for Solid Rocket Motor Nozzles, Energies, 9 (2016) 430-450

[4] Lin Sun, Futing Bao, Yu Zhao, Lian Hou, Ning Zhang, Weihua Hui and Wei Shi, Crack Cause Analysis of a Graphite Nozzle Throat Insert, Acta Astronautica, http://dx doi org/10 1016/j actaastro 2017 04 016

[5] Nigar B , Analysis of Crack Initiation and Propagation in Solid Rocket Motor Nozzle Throats, Middle East Technical University, Master's Degree Thesis, November 2019

[6] Solid Rocket Motor Nozzles NASA SP-8115, 126 pages, published by NASA, Washington, D C , 1975

[7] Shields, S , Scout Nozzle Data Book, NASA, 1976

[8] Unlu D , Mumcu B , Olgar T , Füze-Roket sistemlerinde kullanılan lüle çekirdeği termomekanik analizi ve tasarım iyileştirmesi, VII Ulusal Havacılık ve Uzay Konferansı, 2018

Investigation of Effects of Wavy Leading Edge on Flow Around A Naca 0020 Airfoil

M.SEYHAN[1], M. SARIOĞLU[2] and Y.E. AKANSU[3]

[1] Karadeniz Technical University, Trabzon/Turkey, mehmetseyhan@ktu.edu.tr
[2] Karadeniz Technical University, Trabzon/Turkey, sarioglu@ktu.edu.tr
[3] Niğde Ömer Halisdemir University, Niğde/Turkey, akansu@ohu.edu.tr

Abstract **- An experimental study is carried out to investigate the effects of leading-edge tubercles on NACA 0020 airfoil at Re = 63000. Force measurements are carried out with the help of a load cell at different angles of attack between 0-30 degrees. Tubercle parameters are selected as a = 0.05c and λ = 0.5c for M1 and a = 0.025c and λ = 0.5c for M2. M2 and M1 tubercle models produced 143% and 110% more lift than the baseline model at α = 10°, respectively. They also postponed the stall with the help of tubercle geometry and change the common stall characteristics of the baseline model. The airfoil with leading-edge tubercle showed better lift performance as compared to the baseline airfoil at the post-stall region.**

Keywords **– NACA 0020, Lift coefficient, Drag coefficient and Tubercle.**

I. Introduction

In recent years, there is a growing interest in biomimicry so as to use in flow control applications for the purpose of enhanced the aerodynamic characteristics such as lift, drag, vortex shedding etc. A pectoral fin of the humpback whale, having giant body size and good maneuverability skill has drawn attention by the researchers. It was the first time Fish and Battle [1] have proposed to investigate the flipper of the humpback whale. They also found that the flipper of the mid-span section is familiar to NACA 63_4-021 airfoil profile. Thereafter, Miklosovic et al. [2] investigated the aerodynamic performance of humpback whale flipper with NACA 0020 airfoil section at the Reynolds number ranges between $5.05x10^5$ and $5.2x10^5$. Their results indicated that NACA 0020 airfoil with the leading-edge tubercle has a substantial effect on aerodynamic characteristics such as lift coefficient and stall behavior. In the latter study of Miklosovic et al. [3], they try to find stall performance enhancement originated rectangular airfoil with leading-edge tubercle or semi-span airfoil with tubercle. They revealed that the rectangular airfoil leads to a decrease in lift and increase in drag coefficient but semi-span airfoil lead to an increase in lift and moment and a decrease in the drag coefficient. Johari et al. [4] researched the effects of wavy leading edge shape parameters such as the amplitude and wavelength. According to mean chord, these parameters were chosen as 0.12c, 0.05c and 0.025c for the amplitude (a) and 0.5c and 0.25c for the wavelength (λ). They performed tufts flow visualization and force measurement experiments. Their results indicated that leading-edge tubercle leads to a decrease in lift and increase in drag at the pre-stall region. On the other hand, it leads to an increase in lift and delay the stall at the post-stall region. In addition, they found that an optimum improvement for lift coefficient, drag coefficient and stall behavior optimum amplitude obtain with small amplitude (0.025c) and high wavelength (0.5c). In their results, this tubercle geometry did not present the common stall characteristics as compared to baseline airfoil. Pedro and Kobayashi [5] performed numerical study to investigate the humpback whale flipper having NACA 0020 airfoil section at Re = 50000. The flipper model has large streamwise vortices in contrast to the flipper model without leading edge tubercle. In the midsection of the flipper, high vorticity lead to enhance the boundary layer by moving high momentum flow near to the airfoil surface. They suggested that leading edge tubercle flow structure is similar to a wing fence that is flow separation control device. Hansen et al. [6] investigated the effect of NACA 0021 and NACA 65-021 airfoils with leading-edge tubercle at $Re = 1.2x10^5$. Force measurement and hydrogen bubble flow visualization are performed. The leading-edge parameters are selected as 0.03c - 0.11c for the amplitude and 0.11c – 0.86c for wavelength. They observed that higher amplitude and lower wavelength lead to higher lift coefficient and lower drag coefficient, especially at the post-stall region. They suggested that the working mechanism of the leading-edge tubercle is familiar with the vortex generator. Weber et al. [7] carried out to numerical study by using STAR-CCM and Solidworks flow simulation and their results were compared with the experimental results. Their results indicated that the lift coefficient has a good fit at the post-stall region but the drag coefficient has not a good fit at this region. Proper orthogonal decomposition (POD) method and particle image velocimetry (PIV) flow visualization methods are applied to investigate the flow structure around NACA 634-021, having the tubercle leading edge by Wei et al. [8]. Their results indicated that the tubercle in leading-edge cause to rise flow instability due to vorticity structure. While flow separation and reattachments are prominent in the middle and trough of the tubercle, turbulent disturbance is prominent in the peak of the tubercle. Corsini et al. [9] numerically

investigated the effect of leading-edge tubercle on the flow around the NACA 0015 airfoil. Tubercle model has enhanced the post-stall characteristic of the model.
The aim of this study is to research the influence of the NACA 0020 airfoil with and without the tubercle leading edge in the suction type wind tunnel. To do this, force measurement is carried out by using a load cell at Re = 63000.

II. Experimental Setup

Experiments were implemented in the suction type low-speed wind tunnel located in Karadeniz Technical University Mechanical Engineering Department. This tunnel has a test section of 57 cm x 57 cm cross-section. Free stream velocity of the tunnel can adjustable up to 50 m/s by using frequency inverter. NACA 0020 airfoil model having 150 mm chord and 450 mm spanwise length was employed in this study. Leading-edge tubercle models are also used as NACA 0020 section. The test models were manufactured by using CCH Technology brand MY Z35 model 3D printer from PLA filament. The test models were equally divided into 2 parts (each of 225 mm). Manufactured models were carefully assembled and sanded with sandpaper from coarse to fine. As shown in Figure 1(b), the experimental setup is composed of a load cell, a rotary unite, a rod, two endplates and a rod protection ring. The test model is rotated with rotary unite at the angle of attacks between 0° and 30° with an increment of 1°. There is a gap between end plate and the test model for the force measurement. The test model is vertically placed in the tunnel.

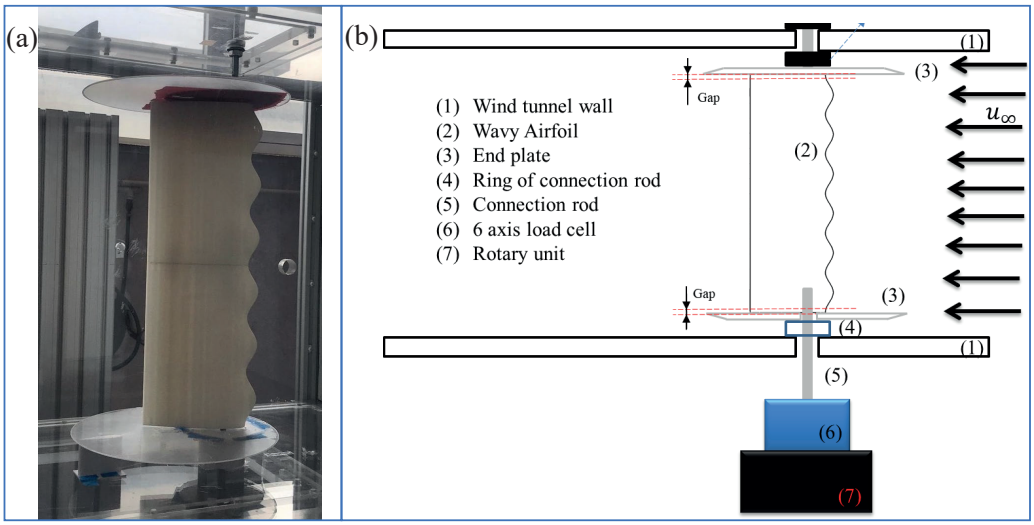


Figure 1: (a) Leading edge tubercle NACA 0020 airfoil in the test section and (b) schematic view of the force measurement setup.

In this study, tubercle parameters were selected as amplitude (a) of 0.025c and 0.05c, and wavelength (λ) of 0.5c as indicated by Johari et al. [4]. Schematic view of the baseline and tubercle models are shown in Fig. 2. M1 denotes the tubercle model having a = 0.05c and λ = 0.5c. M2 denotes the tubercle model having a = 0.025c and λ = 0.5c. Figure 1(a) shows the manufactured M1 model in the test section of the wind tunnel.

Tubercle curve for NACA 0020 airfoil is defined in Equation (1).

$$F(x)=(a*\cos((2\pi x)/\lambda)) \quad (1)$$

6 axis ATI GAMA load cell is used in the force measurement experiments. Force data are taken as 10000 samples during 20 seconds. These experiments were performed at Reynolds number of 63000 based on the mean chord of 150 mm for the tubercle and baseline models.

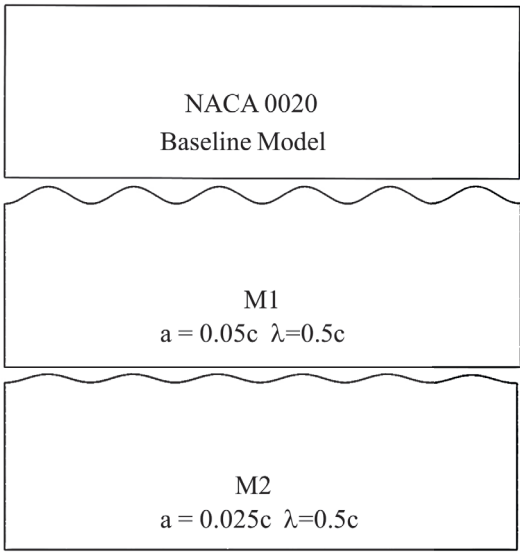


Figure 2: Schematic view of the test models.

III. RESULTS

The force measurements including lift and drag are performed to investigate the influence of the leading-edge tubercle on the flow around NACA 0015 airfoil at Re = 63000. The lift coefficient (C_L) variation for the baseline model and leading-edge tubercle models that are M1 and M2 are plotted as a function of angle of attack (AOA) between 0° and 30° in Figure 3. Stall angle of the baseline model is α = 7°. Up to this angle, C_L of the baseline model is slightly higher than that of M1 and M2 models. After the stall angle, tubercle M2 delayed stall angle from α = 7° to α = 10° and increased the lift coefficient when compared to the baseline model. M2 and M1 tubercle models produce 143% and 110% more lift than the baseline model at α = 10°, respectively. In the post-stall region, M2 and M1 produce higher lift than the baseline model from α = 8° up to α = 18° and α = 23°, respectively. Lift performance of M2 is better than that of M1 and the baseline model due to increasing the maximum lift and delaying the stall angle. M1 has postponed the stall angle from α = 7° to α = 9°. At the post-stall region, C_L of M1 changes in a narrow band that is C_L =0.52 - 0.63. M2 provides 16% more lift coefficient than M1 at α = 10°.

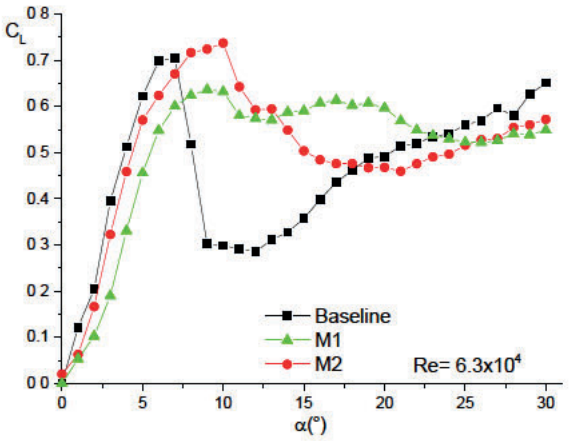


Figure 3: The variation of the lift coefficient as a function of angle of attack for Baseline, M1 and M2 models.

The drag coefficient (C_D) variations of the baseline, M1 and M2 tubercle models as a function of angle of attack can be seen in Figure 4. The baseline and tubercle models show similar trend up to AOA = 7° that is stall angle. After this angle, there is a sudden jump in the drag coefficient of the baseline because of the stall. In the stall situation for the baseline, there is a sudden decrease in lift coefficient and increase in the drag coefficient while there is no significant change in drag and lift coefficient of the wavy models (M1 and M2) due to absence of common stall characteristic. While M1 and M2 tubercle models show familiar drag coefficient trend almost all angle of attacks, The most prominent difference between baseline model and tubercle models (M1 and M2) has been seen at the AOA ranges between 7°-11° and 19°-30°

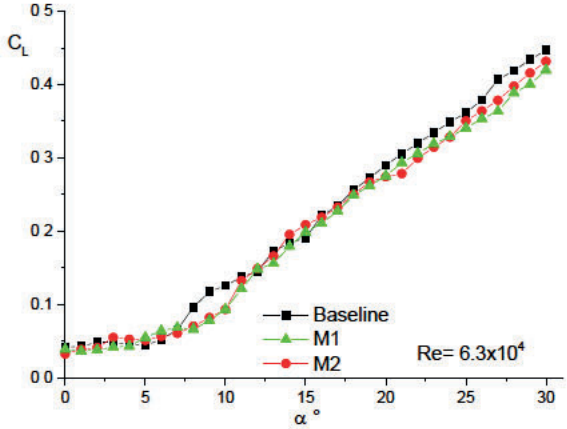


Figure 4: The variation of the drag coefficient as a function of angle of attack for Baseline, M1 and M2 models.

Lift to drag ratios (L/D) are expressed as airfoil efficiency. Figure 5 shows the lift to drag ratios as a function of the angle of attack for the baseline, M1 and M2 models. L/D ratios of the baseline have a similar trend with M2 up to AOA= 6°. At post-stall region, M1 and M2 have greater L/D ratios than the baseline model.

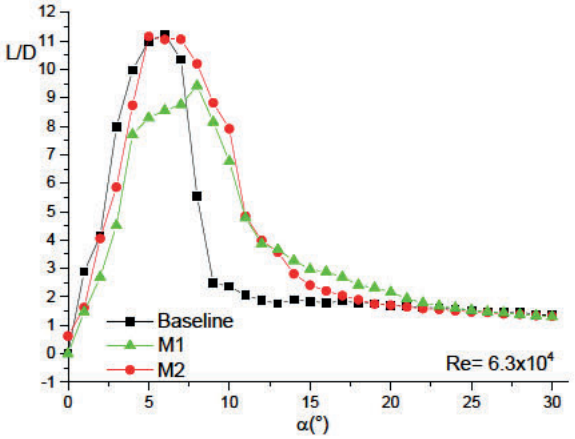


Figure 5: The variation of lift to drag ratios as a function of angle of attack for Baseline, M1 and M2 models.

Normalized lift (L_N) and Normalized drag (D_N) is the ratio of lift or drag coefficient of the wavy model to that of the baseline model. If L_N is bigger than 1, this indicates airfoil lift performance improvement. If D_N is lower than 1, this indicates drag coefficient improvement according to the baseline model. Figure 6 (a and b) shows the variation of the normalized lift and drag coefficient by the baseline model, respectively. While L_N of M1 and M2 is lower than 1 up to AOA = 7°, they are higher than 1 at the angle of attack ranges 7°-18° for M2 and 7°-23° for M1. D_N of tubercle M1 and M2 models are lower than 1 almost all angle of attacks except for AOA= 3, 4, 14 and 15 for M2 and 6, 7, 14 and 15 for M1. This means that there is a drag reduction in M1 and M2 as compared to the baseline models.

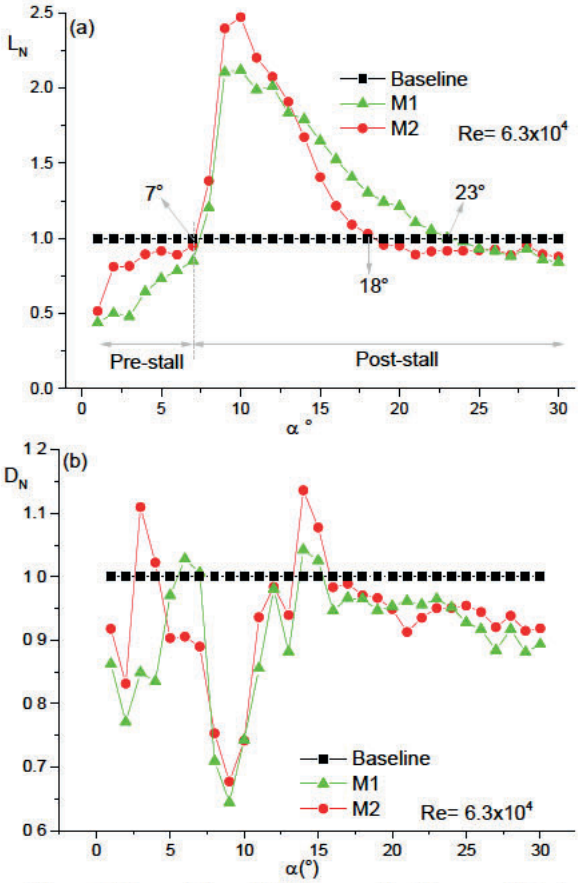


Figure 6: The variation of (a) the normalized drag coefficient and (b) the normalized lift coefficient by the baseline model.

IV. Conclusion

Aerodynamic characteristics of the NACA 0020 airfoil model with and without leading-edge tubercle are experimentally investigated in the suction type wind tunnel at Re=63000. Force measurement experiment is performed by using 6 axis load cell. The results indicate that leading-edge tubercle enhances the lift characteristics at the post-stall region and delay the stall from $\alpha = 7°$ to $\alpha = 10°$. Moreover, the drag coefficient of the tubercle models is lower than the baseline model at almost all AOAs.

Acknowledgment

The authors would like to acknowledge the financial support of this work by the Scientific and Technological Research Council of Turkey (TUBITAK) under the Contract Number of 118M592.

References

[1] F E Fish and J M Battle, "Hydrodynamic-Design of the Humpback Whale Flipper," *Journal of Morphology*, vol 225, no 1 pp 51–60, 1995, doi: 10 1002/jmor 1052250105

[2] D S Miklosovic, M M Murray, L E Howle, and F E Fish, "Leading-edge tubercles delay stall on humpback whale (Megaptera novaeangliae) flippers," *Phys. Fluids*, vol 16, no 5, pp 39–42, 2004, doi: 10 1063/1 1688341

[3] D S Miklosovic, M M Murray, and L E Howle, "Experimental Evaluation of Sinusoidal Leading Edges," *J. Aircr.*, vol 44, no 4, pp 1404–1408, 2007, doi: 10 2514/1 30303

[4] H Johari, C Henoch, D Custodio, and A Levshin, "Effects of leading-edge protuberances on airfoil performance," *AIAA J.*, vol 45, no 11, pp 2634–2642, 2007, doi: Doi 10 2514/1 28497

[5] H Carreira Pedro and M Kobayashi, "Numerical Study of Stall Delay on Humpback Whale Flippers," *46th AIAA Aerosp. Sci. Meet. Exhib.*, no January, pp 7–10, 2008, doi: 10 2514/6 2008-584

[6] K L Hansen, R M Kelso, and B B Dally, "Performance Variations of Leading-Edge Tubercles for Distinct Airfoil Profiles," *AIAA J.*, vol 49, no 1, pp 185–194, 2011, doi: 10 2514/1 J050631

[7] P W Weber, L E Howle, M M Murray, and D S Miklosovic, "Computational evaluation of the performance of lifting surfaces with leading-edge protuberances," *J. Aircr.*, vol 48, no 2, pp 591–600, 2011, doi: 10 2514/1 C031163

[8] Z Wei, B Zang, T H New, and Y D Cui, "A proper orthogonal decomposition study on the unsteady flow behaviour of a hydrofoil with leading-edge tubercles," *Ocean Eng.*, vol 121, pp 356–368, 2016, doi: 10 1016/j oceaneng 2016 05 043

[9] A Corsini, G Delibra, and A G Sheard, "On the role of leading-edge bumps in the control of stall onset in axial fan blades," *J. Fluids Eng. Asme*, vol 135, no 8, p 81104, 2013, doi: Artn 081104\rDoi 10 1115/1 4024115

[10] D F Kurtulus, "Ability to forecast unsteady aerodynamic forces of flapping airfoils by artificial neural network," *Neural Comput. Appl.*, vol 18, no 4, pp 359–368, 2009, doi: 10 1007/s00521-008-0186-2

Mechanical Analysis of Nano Boron Nitride Doped Adhesive Joints

K. GÜLTEKİN[1], G. UĞUZ[1] and A. ÖZEL[2]
[1]Ondokuz Mayıs University, Samsun/Turkey, kursat.gultekin@omu.edu.tr
[2]Binali Yıldırım University, Erzinca/Turkey, adnan.ozel@erzincan.edu.tr

Abstract **- The new generation nano-composite materials enhanced with using nanoparticle has started been used in many areas due to increase in mercantile value of the material and it's versatile and long-lived texture. Boron-containing products follows an increasing course in the space and aviation industry by increasing up to the durability of the materials by extending the lifetime of the materials. The boron nitride (BN) nano particles called white graphene has superior chemical, electrical and thermal properties. In this research; Mechanical properties of adhesively bonded joints were investigated using two different adhesive material modified with hexagonal BN nanoparticles. In this context; nano adhesives have been produced by adding BN nanoparticles in different proportions (%0.5, %1, %2, %3, %4 and %5) in adhesives (Araldite 2011 and Hexion MGS-L285). With these adhesives, single lap joints aluminum AA2024-T3 were produced according to the ASTM D1002 standard and their mechanical properties under tensile load were examined experimentally. It has been observed that with addition of the BN into the adhesive material, the damage load of the bonding joints increases depending on the contribution rate. Also; distribution of the boron nitride nanoparticles in the adhesive material was investigated with using a scanning electron microscope (SEM-EDS).**

Keywords **- Boron Nitride, Epoxy, Nanoparticle, Polymer, Adhesive Bonding Joints**

I. INTRODUCTION

THE reason for the increase of intensive research on nanomaterials is that nanoparticles are have superior properties and functions in a certain size regardless of their volumetric structure. In order to improve certain properties in polymer and metal materials, graphene, carbon nanotube, fullerene, carbon nanofiber, aluminum oxide, clay, etc. are often preferred [1-3]. Boron nitride (BN), called white graphene, has excellent thermal, electrical, and chemical properties. Having these properties has caused to find a wide application area compared to carbon based and composites boron nitride doped polymers (graphene, carbon nanotube etc.) [4-7].

When the studies in the literature are examined; It is seen that there are improvements in the mechanical and thermal properties of composite materials produced by doping BN nanoparticles. Jakubinek et al. stated that boron nitride nanotubes (BNNT) would interact better with epoxy than carbon nanotubes. In this case, BNNTs are specified as nanotubes that can be preferred to strengthen epoxy, especially in properties such as transparency / color, neutron absorption or electrical insulation, when electrical conductivity is not required [8]. Chen et al. [9] stated that polymer materials and boron nitride nanotubes form a stronger adhesion interface than carbon nanotubes. It was determined that the BNNT-epoxy interface strength was higher than the BNNT-Polymethyl methacrylate interface strength. The extraordinary load transfer capacity findings of BNNT polymer interfaces show that BNNTs are excellent reinforcing nanofiller materials for lightweight and high strength polymer nanocomposites.

In a study by Rahmat et al. they observed an increase of up to 7% in the tensile strength of composite materials with BNNT, an increase of up to 16% in the modulus of elasticity, and a decrease of 9% in the fracture deformation [10].

Navaneethakrishnan et al. [11] in their study on BN / epoxy nanocomposites containing different proportions of BN particles; while the tensile strength, bending strength, impact resistance, micro hardness values of nanocomposites increased up to 1% by mass of BN, these values decreased after 1% by mass. The lowest wear amount and friction coefficient were observed in BN-epoxy nanocomposite containing 1% BN. It has been stated that the wear rate and coefficient of friction of nanocomposites decrease due to the load bearing properties of BN particles. With these adhesives, single lap joints of aluminum AA2024-T3 alloys were produced according to the ASTM D1002 standard and their mechanical properties under tensile load were examined experimentally. Additionaly, the distribution of boron nitride nanoparticles in the adhesive was investigated by using SEM-EDS.

In the presented study; the mechanical properties of bonding joints were investigated using two different adhesives modified with hexagonal BN nanoparticles. In this context; Nanoadhesives were produced by adding different proportions (0.5%, 1%, 2%, 3%, 4% and 5%) of BN nanoparticles into adhesives (Araldite 2011 and Hexion MGS-L285).

II. MATERIALS AND METHODS

In the study; AA2024-T3 aluminum alloy is used as the bonded material to produce single lap joints. Aluminum is one of the most used materials after steel and cast iron in the field of engineering. Especially, AA2024 alloy has low density (2.78 gr / cm^3), high strength (E = 73.1 GPa, σt = 485 MPa, σak = 345 MPA), electrical conductivity (30% IACS) and corrosion resistance. So, it is a material that is increasingly used in automotive and aerospace industry.

As an adhesive; Araldite-2011 (Hunstman, USA) and MGS-L285 (Hexion, USA) two-component liquid epoxy adhesives were used. Araldite 2011 epoxy adhesive; It has an epoxy / hardener composition ratio of 1: 0.8 by weight and the curing process is completed in 75 minutes at 60 ° C. MGS-L285 epoxy adhesive is; It has an epoxy / hardener composition ratio of 1: 0.4 by weight and cures at 80 ° C in 50 minutes.

Nanoparticles as used for modifying the adhesives are approximately 120 nm and hexagonal boron nitride having 97% purity (h-BN) nanoparticle (Bortek Inc., Turkey). The SEM image of the h-BN nanoparticle is given in Figure 1.

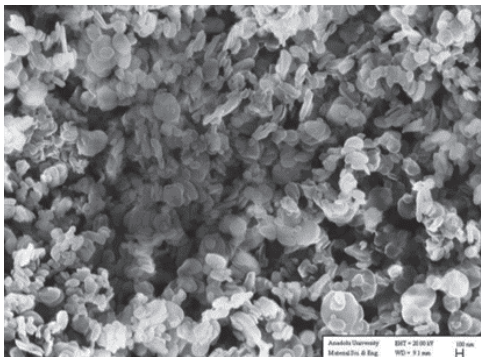


Figure 1: SEM image of h-BN particle.

In this study, the mechanical properties of the h-BN modified single lap joint samples were experimentally investigated. Experimental parameters and joint geometry produced in accordance with ASTM D1002 standard are given in Table 1 and Figure 2.

Table 1: Parameters of experiments.

Sample	Adhesive	Doping ratio %
A0	Araldite 2011	0
ABN0.5	Araldite 2011	0.5
ABN1	Araldite 2011	1
ABN2	Araldite 2011	2
ABN3	Araldite 2011	3
ABN4	Araldite 2011	4
ABN5	Araldite 2011	5
M0	MGS-L285	0
MBN0.5	MGS-L285	0.5
MBN1	MGS-L285	1
MBN2	MGS-L285	2
MBN3	MGS-L285	3
MBN4	MGS-L285	4
MBN5	MGS-L285	5

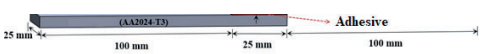


Figure 2: Single lap joint geometry and dimension.

Before the production step of the hexagonal boron nitride nanoparticles doped and undoped single lap joint samples shown in Figure 2, doped and undoped adhesives have been prepared. At this step; h-BN nanoparticles (in a ratio appropriate to the weight of the epoxy to be made were weighed and mixed with acetone in an ultrasonic homogenizer for 15 minutes in a beaker. Thus, the separation of clustered h-BN particles was achieved. Epoxy was added to the h-BN / acetone solution and mixed in an ultrasonic homogenizer for 30 minutes. The mixing process was carried out in an ice bath so that the heat released during mixing does not damage the nanoparticle and epoxy structures. In order to evaporate the acetone in the solution, the solution was kept in the oven. After the acetone has evaporated, adhesives with additives are produced by adding an appropriate amount of hardener to the epoxy + h-BN mixture. Also for additive-free adhesives; acetone was added and mixed in an ultrasonic homogenization for 30 minutes and then the same steps were followed (Figure 3).

Figure 3: Production of adhesives.

After AA2024 aluminum alloy sheets were cut in the dimensions specified in Figure 2, the burrs arising from the cutting process were removed with sandpaper. Oil, dirt, etc. formed on the sample surface during the production and cutting of the aluminum plate were pre-cleaned with acetone. The aluminum alloys, whose surface was cleaned, were then mechanically smoothed with 1200 SiC sandpaper. The strength of the bonding joints depends on the flawless surface. After mechanical cleaning, samples were etched with sulfuric acid / sodium dichromate solution for chemical surface treatment. For this purpose; aluminum samples were etched into the prepared (310 g/L sulfuric acid, 67.3 g/L sodium dichromate) sulfuric acid/sodium dichromate etching solution) at 65-70 ℃ for 15 minutes. The etched samples were washed with pure water and dried (Figure 4).

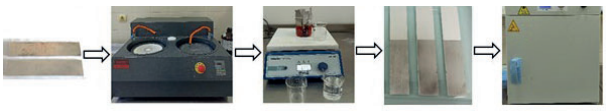
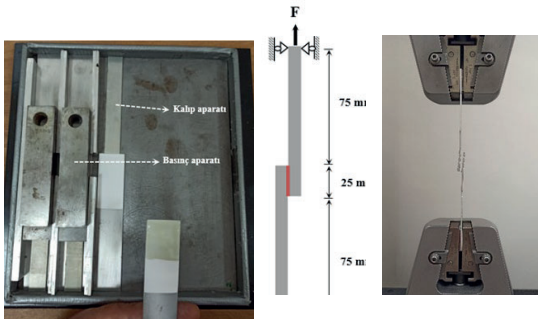


Figure 4: Surface preparation steps of AA2024-T3 alloy.

After the preparation of the adhesives and adherend materials was completed, AA2024-T3 joint samples were produced using the mold shown in Figure 5. Mold and pressure apparatus are used to adjust the uniform adhesive thickness (Figure 5.a). The bonded samples were then kept in the oven (75 minutes at 60 ℃ for Araldite, 50 minutes at 80 ℃ for MGS-L285) and cured.

After the curing step of the samples, excess adhesives were removed and the production of the connection samples was completed. The tensile test of the produced joint samples was carried out in the INSTRON universal tensile device at the boundary conditions shown in Figure 5.b at a tensile speed of 1 mm/min.



(a) (b)

Figure 5: Production of single lap joint samples and tensile testing.

III. CONCLUSIONS

The average failure loads obtained from the tensile test of single AA2024-T3 joint samples with and without hexagonal boron nitride are given in Figure 6 and Figure 7.

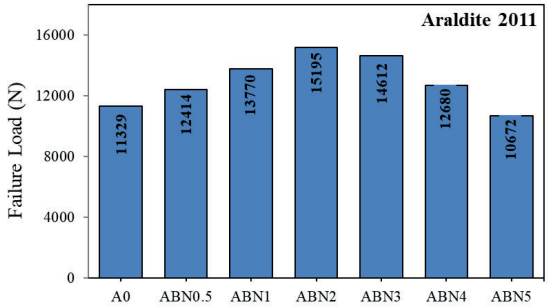


Figure 6: Average failure loads for Araldite 2011 adhesive.

Adding boron nitride nanoparticles to Araldite 2011 adhesive; it is seen from experimental results that an increase in the failure load of up to 4% doped were shown on single lap joints (Figure 6). With the addition of 0.5% BN into the adhesive, the failure load of the bonding joint increases by 11% and by 21% with the doping of 1% BN. While the biggest increase in the failure load occurred at the rate of 2% BN, it is seen that after this rate (3%, 4% and 5%), the failure load decreased. With 5% BN reinforcement in the adhesive, the failure load of the joints is reduced by 6%. The reason for this can be explained as the occurrence of agglomerations with the increase of BN particles in the epoxy. This state is also seen when SEM-EDS images are examined (Figure 8 and Figure 9). As the agglomeration of BN particles occurring in the epoxy resin creates a notch effect, it causes rapid crack propagation. In other additive rates, it causes an increase in the failure load of the joints as the BN particles are dispersed homogeneously in the epoxy. In the boron nitride particle, the nitrogen atoms contain more protons than the boron atoms, causing more electrons in their bond structure. The irregular bond structure of epoxies creates active bond structures between boron nitride and epoxy, resulting in stronger bond formation.

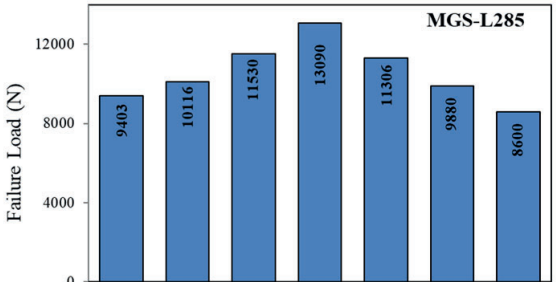


Figure 7: Average failure loads for MGS L285 adhesive.

When the average failure loads of the joints obtained by adding BN particles into MGS-L285 epoxy resin in different proportions are examined; it is seen that the failure load increases up to a certain rate depending on the BN particles additive ratio. The average failure load of the joint obtained by adding 2% BN into the adhesive has increased by approximately 40% compared to the undoped joint. In the joints obtained with 5% BN reinforcement, as a result of the agglomeration created by the BN particles in the epoxy, the failure load was reduced by approximately 9%. Failure loads increased by approximately 8%, 22%, 20% and 5% for 0.5%, 1% and 3% and 4% with BN addition ratio respectively. It has been observed that with the addition of BN into the adhesive, the failure load of the joints increases depending on the rate of addition.

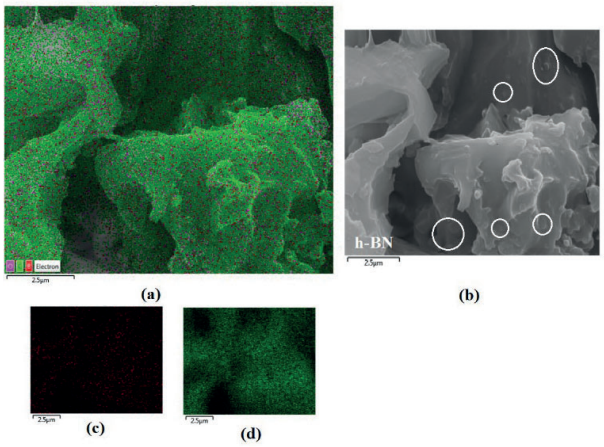


Figure 8: SEM-EDS image of Araldite 2011 adhesive with 2% h-BN additive; a)EDS, b)SEM, c)Boron d)Carbon.

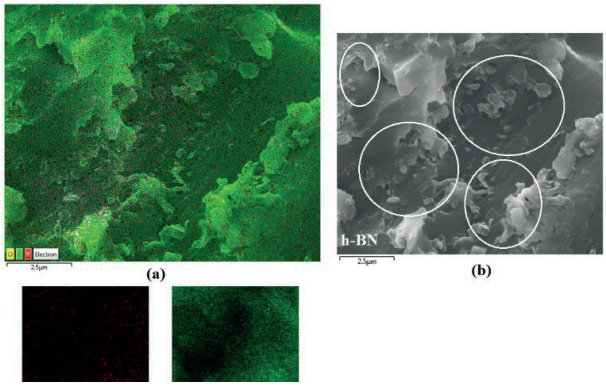


Figure 9: SEM-EDS image of Araldite 2011 adhesive with 5% h-BN additive; a)EDS, b)SEM, c)Boron d)Carbon.

In order to determine the distribution of BN particles in Araldite 2011 adhesive, 2% and 5% doped bulk samples were analyzed using the SEM-EDS technique. It was observed that the particles of Araldite 2011 were partially dispersed into the epoxy resin with the addition of 2% BN, and as the ratio increased (5%), BN nanoparticles agglomerated partially.

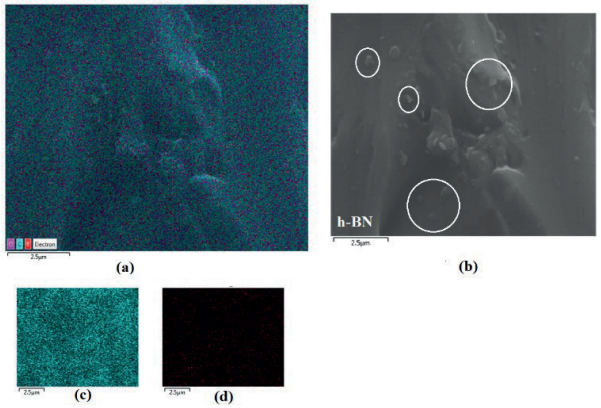


Figure 10: SEM-EDS image of MGS L285 adhesive with 2% h-BN additive; a)EDS, b)SEM, c)Boron d)Carbon.

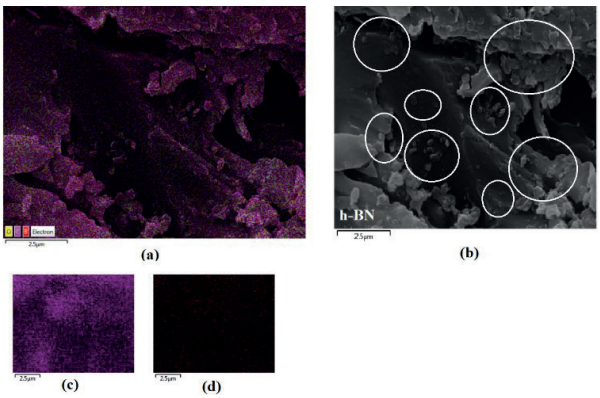


Figure 11: SEM-EDS image of MGS L285 adhesive with 5% h-BN additive; a)EDS, b)SEM, c)Boron d)Carbon.

In order to determine the distribution of BN particles in MGS-L285 adhesive, 2% and 5% BN added adhesive bulk samples were analyzed using SEM-EDS technique. It was observed that the particles were partially dispersed in the adhesive with the addition of 2% BN, and as the ratio increased (5%), the aggregation of BN nanoparticles was observed (Figure 10 and Figure 11).

ACKNOWLEDGMENT

This study was supported by the Scientific and Technological Research Council of Turkey (TUBITAK) with 119M939 project number.

REFERENCES

[1] K Gültekin, S Akpınar, A Gürses,, Z Eroğlu, S Çam, H Akbulut, Z Keskin, A Özel , "The effects of graphene nanostructure reinforcement on the adhesive method and the graphene reinforcement ratio on the failure load in adhesively bonded joints," *Composites Part B: Engineering,* vol 98, pp 362-369,2016

[2] P Jojibabu, Y X Zhang, and P Gangadhar, "A review of research advances in epoxy-based nanocomposites as adhesive materials," *International Journal of Adhesion and Adhesives*, vol 96,102454,2020

[3] P Jojibabu, Y X Zhang, A N Rider, J Wang, R Wuhrer, P Gangadhar, "High-performance epoxy-based adhesives modified with functionalized graphene nanoplatelets and triblock copolymers," *International Journal of Adhesion and Adhesives*, vol 98,10521,2020

[4] X Huang, C Zhi, P Jiang, D Golberg, Y Bando, Y Tanaka, "Polyhedral Oligosilsesquioxane-Modified Boron Nitride Nanotube Based Epoxy Nanocomposites: An Ideal Dielectric Material with High Thermal Conductivity," *Advance Functional Materials*, vol 23,pp 1824-1831,2013

[5] K S Kim, M J Kim, C Park, C C Fay, S Chu, C T Kingston, B Simard, "Scalable manufacturing of boron nitride nanotubes and their assemblies: a review," *Semiconductor. Science and Technology,* vol 32, 013003,2017

[6] N O Bansal, J B Hurst, and S R Choi, "Boron nitride nanotube-reinforced glass composites,"*Journal of American Ceramic Society,*vol 89, pp 388–390, 2006

[7] W Meng, Y Huang, Y Fu, Z Wang, and C Zhi, "Polymer composites of boron nitride nanotubes and nanosheets," *Journal of Material Chemistry C*, vol 2, pp 10049–10061, 2014

[8] M B Jakubinek, B Ashrafi, Y Martinez, M Rahmat, M Yourdkani, K S Kim, K Laqua, A Yousefpor, B Simard, "Nanoreinforced epoxy and adhesive joints incorporating boron nitride Nanotubes,"*International Journal of Adhesion and Adhesives,* vol 84,pp 194–201,2018

[9] X Chen, L Zhang, C Park, C C Fay, X Wang, C Ke,"Mechanical strength of boron nitride nanotube-polymer interfaces," *Applied Physics. Letter*, vol 107,253105,2015

[10] M Rahmat, A Naftel, B Ashrafi, M B Jakubinek, Y M Rub, B Simard, "Dynamic Mechanical Characterization of Boron Nitride Nanotube Epoxy Nanocomposites," *Polymer Composites Society of Plastics Engineers,* vol 40,pp 2109-2131,2018

[11] G Navaneethakrishnan, T Karthikeyan, S Saravanan, V Selvam, "Influence of boron nitride on morphological, mechanical, thermal and wear characteristics of epoxy nanocomposites," *Materials Research Innovations*, vol 24, pp 257-262, 2020

[12] K Gültekin, S Akpınar, A Özel, "The effect of the adherend width on the strength of adhesively bonded single lap joint Experimental and numerical analysis," *Composites Part B: Engineering*, vol 60,pp 736-745, 2014

Investigation of Structural Properties of Boron Carbide Reinforced Epoxy Adhesives

G. UĞUZ[1] and K. GÜLTEKİN[2]
[1]Ondokuz Mayıs University, Samsun/Turkey, gediz.uguz@omu.edu.tr
[2]Ondokuz Mayıs University, Samsun/Turkey, kursat.gultekin@omu.edu.tr

Abstract **- Research, conducted in recent years, has shown that boron compounds, in particular, can be an important type of reinforcement in improving the mechanical and thermal properties of polymer materials. Within the scope of this study presented, the effect of nanoparticle reinforcement ratio in adhesives, new adhesive samples were prepared by adding different proportions (0%, 0.5%, 1%, 2%, 3%, 4% and 5%) of nano hexagonal boron nitride particles into two different epoxy adhesives. The prepared samples were characterized by Fourier Transform Infrared Spectroscopy (FT-IR) and Scanning Electron Microscopy Energy Disperison X-Ray Sepctroscopy (SEM-EDS) techniques. SEM-EDS images show that the nano-reinforcement material is dispersed in the polymer matrix, but as the reinforcement ratio increases, it cannot sufficiently adhere to the polymer network. Structural analysis based on FT-IR spectra show that the permeability values decrease in peak intensities due to the increase in the reinforcement ratio. As a result of the obtained findings, it is recommended to make modifications to nano h-B4C particles for adhering to polymer network**

Keywords **- Adhesive, Boron Carbide, Epoxy, Nanoparticle, SEM-EDS.**

I. INTRODUCTION

THE functionality of materials, which are the basic inputs of industrial fields such as micro-electronics, biotechnology and nanotechnology, has a great importance in our lives. In line with this importance, various studies are carried out by researchers to investigate and develop materials. There are several requirements for the production of advanced materials, especially for the defense, aerospace, micro-electronics, communications and automotive industries.In order to meet these requirements, a multi-disciplinary and process-intensive field is required [1].

Polymer nanocomposites, which have a wide range of applications from structural and coating materials to medical products and electronic devices, are two-component systems produced by mixing polymer matrix and nanoparticle in industry [2,3]. Epoxy resins are widely used in polymer nanocomposite production [4]. Recent research has shown that when nanoparticles are homogeneously dispersed within the epoxy matrix, the resin's mechanical, optical, thermal, electronic, mechanical or magnetic properties are dramatically improved. For this reason, the researchers emphasized that parameters such as size, shape, reinforcement ratio and chemical composition have a significant effect on nanocomposite materials and examined in detail the interaction of nanoparticles in the epoxy matrix [5].

As a result of the researches, it is seen that there are significant improvements in the mechanical, thermal and physical properties of h-B_4C / epoxy composites [6]. It was concluded that h-B4C reinforcement into epoxy could improve mechanical properties such as toughness and modulus of elasticity [5,7].

Two-component epoxy adhesives are produced by polymerization of the oxirane ring, which is the main component in the epoxy resin, with various curing agents. Curing agents are also known as hardeners in two component systems. In addition to the hardener, various additives such as diluents, plasticizers and accelerators are added to these systems for various purposes. Examples of common types of hardeners include aliphatic amines and polyamides, which can generally cure at room temperature. Amine groups are used for curing low molecular weight liquid epoxies [12].

FT-IR is an instrumental technique that examines bond interactions between atoms of materials under infrared radiation. While infrared light is transmitted over the sample during analysis, this beam is absorbed by the sample and creates different vibrations between the bonds in the sample. The remaining infrared ray is reflected. Since each material is made up of different atoms, two different molecules do not give the same spectrum. In this context, FT-IR spectroscopy is a common technique used to determine the components of mixtures, functional groups and structures of unknown materials [13].

In this study, in order to examine the effect of nanoparticle reinforcement ratio in different proportions (0%, 0.5%, 1%, 2%, 3%, 4% and 5%) of nano-hexagonal boron carbide (h-B_4C) were used in two different epoxy adhesives. New adhesive samples produced by the addition of particles were prepared. In order to examine the interaction between epoxy matrix and nanoparticle in the prepared samples, they were characterized by Scanning Electron Microscopy and Energy Dispersion X-Ray Spectroscopy (SEM-EDS) and Fourier Transform Infrared Spectroscopy (FT-IR) and techniques.

II. MATERIALS AND METHODS

As the polymer matrix, two-component Araldite 2011 (Hunstman, USA) adhesive and hardener and MGS-L285 (Hexion, USA) adhesive and hardener were used. The nanoparticle, nano h-B_4C used as reinforcement material was supplied from Nanografi Nano Teknoloji A.Ş. Nano h-B_4C is spherical, its dimensions are between 40-60 nm and it has 99.5% purity and its SEM image and strucutre are shown in Figure 1.

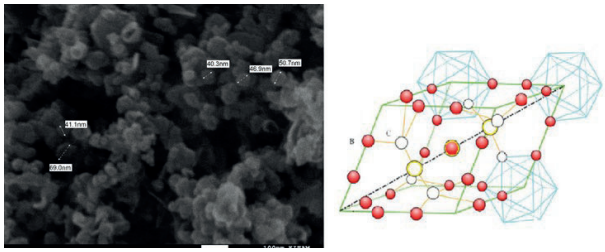


Figure 1: a) SEM image b) structure of h-B_4C particle.

Nano h-B4C particles were reinforced into two different epoxy adhesives in different proportions: 0%, 0.5%, 1%, 2%, 3%, 4% and 5%, and new adhesive samples were prepared. It is important to distribute the nanoparticles homogeneously in the adhesive. For this purpose, in the first stage, nano h-B_4C particles were mixed in approximately 20 ml of acetone for 15 minutes in an ultrasonic homogenizer. In the next step, the nanoparticles dispersed in acetone were mixed with an ultrasonic homogenizer for 30 minutes in the epoxy adhesive, the amount of which was determined according to the mixing ratio. The prepared nanoparticle epoxy mixture, a hardener ratio of 100: 80 in weight of Araldite 2011 epoxy adhesives and a hardener ratio of 100:40 in weight of MGS-L285 epoxy adhesives was added. This mixture is poured into a specially designed and produced female silicone mold. New adhesive samples prepared with Araldite 2011 adhesive poured into silicone mold were cured for 75 minutes at 60 °C and new adhesive samples prepared with MGS-L285 adhesive for 35 minutes at 80 °C in the oven. The preparation steps of the samples are shown in Figure 2 in summary. In order to analyze the samples produced by SEM-EDS and FT-IR techniques, their surfaces were sanded, washed with ethyl alcohol and dried. The contents and codes of the prepared samples are shown in Table 1 and SEM-EDS and FT-IR samples are shown in Figure 3.

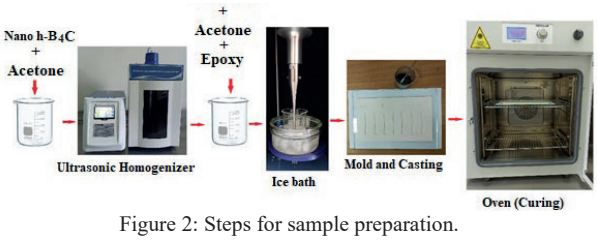


Figure 2: Steps for sample preparation.

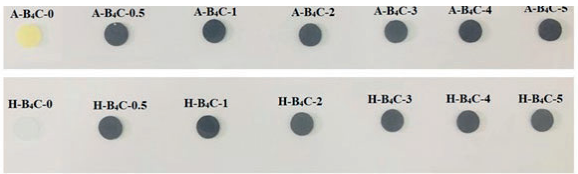


Figure 3: Samples for FT-IR and SEM-EDS.

Table 1: Parameters of experiments.

Sample	Adhesive	Doping ratio %
A-B_4C-0	Araldite 2011	0
A-B_4C-0.5	Araldite 2011	0.5
A-B_4C-1	Araldite 2011	1
A-B_4C-2	Araldite 2011	2
A-B_4C-3	Araldite 2011	3
A-B_4C-4	Araldite 2011	4
A-B_4C-5	Araldite 2011	5
M-B_4C-0	MGS-L285	0
M-B_4C-0.5	MGS-L285	0.5
M-B_4C-1	MGS-L285	1
M-B_4C-2	MGS-L285	2
M-B_4C-3	MGS-L285	3
M-B_4C-4	MGS-L285	4
M-B_4C-5	MGS-L285	5

JEOL JSM-7001F SEM was used to examine the surface morphology of h-B_4C reinforced epoxy adhesives. The surface of the samples produced with h-B_4C was coated with gold (Au) and palladium (Pd) nanoparticles in order to provide electrical conductivity on the surface, prevent storage and examine in detail. Nano h-B_4C particles in the epoxy adhesive were detected by SEM-EDS technique.

Structural analysis of the samples were made using the FT-IR technique with the help of Perkin Elmer Spectrum Two Spectroscopy. Specific bonds of epoxy and h-B4C were characterized by FT-IR analysis. Each spectrum was acquired in the wavelength range 650-4000 cm-1 with a resolution of 4cm − 1 and 4 scans per second. The area where the sample will be placed and the probe were previously cleaned with the help of ethanol. Scan results were evaluated with Spectrum Two software.

Unreinforced epoxy adhesives and reinforced epoxy samples were analyzed using SEM-EDS to determine the distribution of nanoparticles in the epoxy matrix. Since the surface of the samples is polymer and the network structures and nanoparticles in the epoxy matrix can be observed in detail, they are coated with gold nanoparticles by plasma method. The surfaces of the samples were examined. In Figure 4. and Figure 5., 0%, 1% and 5% nano h-B_4C reinforced Araldite and MGS - L285 adhesives and SEM-EDS images are shown.

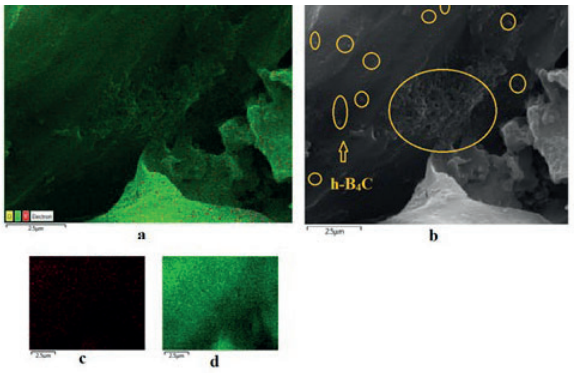


Figure 4: SEM-EDS images of 5% nano h-B_4C reinforced Araldite 2011 epoxy adhesive samples (a)EDS (b)SEM(c)Boron (d) Carbon.

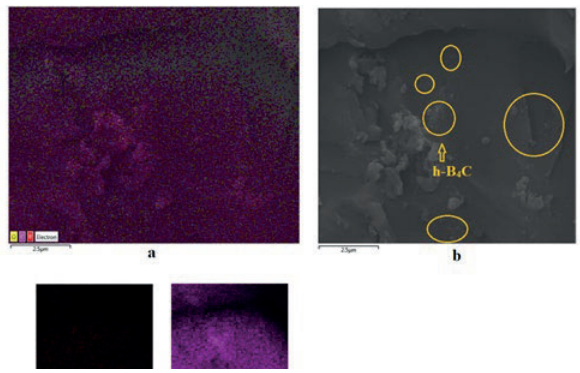


Figure 5: SEM-EDS images of 5% nano h-B4C reinforced Araldite 2011 epoxy adhesive samples a)EDS (b)SEM(c)Boron (d) Carbon.

Within the scope of this study, the structures and curing behavior of new adhesive samples produced by reinforcing h-B_4C particles in different proportions (0%, 0.5%, 1%, 2%, 3%, 4% and 5%) into epoxy adhesive with different properties. It was investigated using the FT-IR technique. FT-IR spectra obtained from the analysis are shown in between Figure 6. and Figure 10. FT-IR peaks are summarized in Table 2.

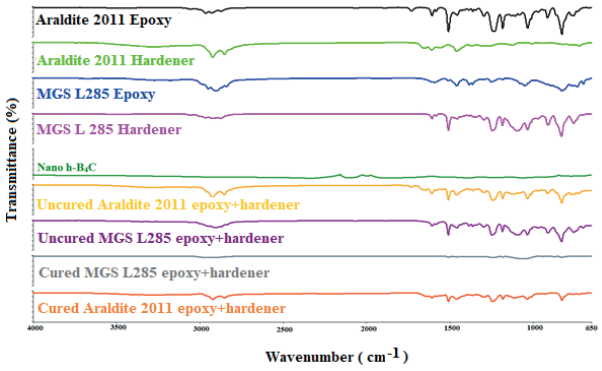


Figure 6: Split FT-IR Spectra of samples.

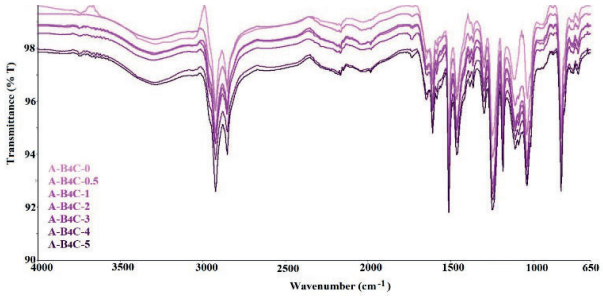


Figure 7: FT-IR spectra of reinforced Araldite 2011 epoxy adhesive samples.

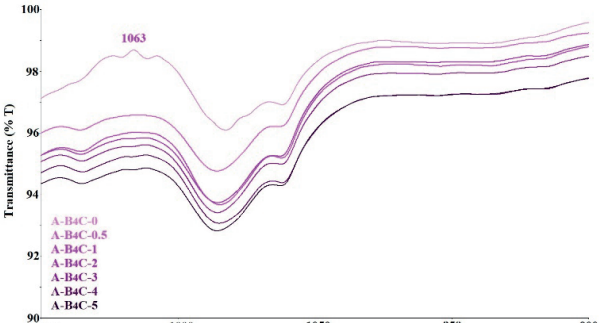


Figure 8: FT-IR spectra of reinforced Araldite 2011 epoxy adhesive samples between 900 -1100 cm^{-1} wavenumber.

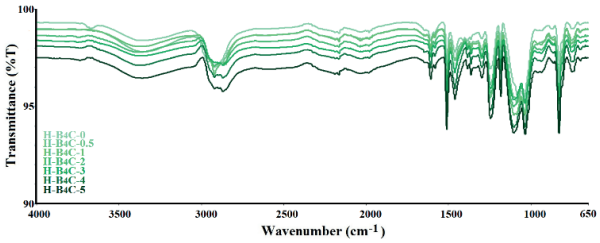


Figure 9: FT-IR spectra of reinforced MGS L285 epoxy adhesive samples.

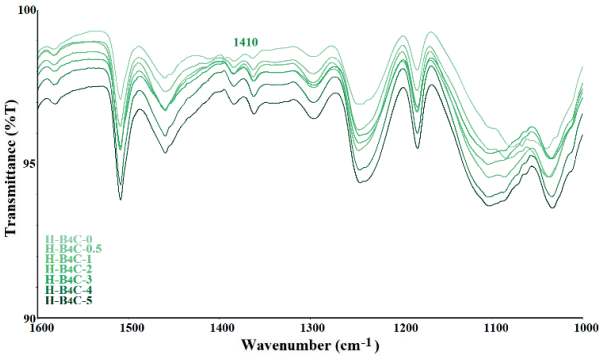


Figure 10: FT-IR spectra of reinforced MGS L285 epoxy adhesive samples between 900 -1100 cm^{-1} wavenumber.

Table 2: FT-IR peaks of samples.

Wavenumber (cm^{-1})	Functional Groups
3381	N-H amine from hardener
3504	O-H stretching (phenol and alcohol)
3057	C-H stretching from oxirane ring
2965	C-H stretching from CH_2
2873	Aromatic C-H
2688	Broad moisture peak
1700	C=O carbonyl peak
1608	Aromatic C=C (ring double bond)
1552	C-B-C bends
1509	Aromatic C-C stretching
1500	Para substituted benzene
1454	Aromatic and aliphatic C-H
1420	B-O stretching vibration
1363	C-CH_3
1260	C-N stretching
1247	Aromatic C-O
1230,	C-O ether bridge
1036	C-O ether
1190	C-O secondary alcohol
1050	B-B vibrations
1035	Aromatic C-O
915	C-O from oxirane ring
915	C-H out of plane bending
840	B-B vibrations
832	B-C bonding
831	C-O-C from oxirane ring
772	CH_2 rocking
701	B-OH
695	C-N from hardener
685	B-B bonding
560	C-C out of plane ring

III. CONCLUSIONS

According to the results obtained with the SEM technique; Unreinforced epoxy samples were found to have a flatter surface than nanoparticle reinforced samples. It has been observed that some of the voids caused by the network structure of the polymer on the surfaces are partially filled with nanoparticles. The unfilled voids seen on the surface indicate that the interaction in the network structure of the nanoparticle and the polymer matrix is not complete [14]. SEM-EDS images obtained from this study show the presence of void regions welded in the network structure of the polymer on the surfaces. The voids seen on the surface indicate that the interaction between the nanoparticle and the polymer matrix is not complete and there are local agglomerations. It is recommended to use the modification methods found in the literature to better clamp the nano h-B_4C surfaces to the mesh. Particularly, while the percentage of nano h-B_4C reinforcement increases from 1% to 5%, partial dispersion is observed.

According to the results obtained from the FT-IR spectra, it has been shown that the nano h-B_4C, which is specified as nano reinforcement in the polymer matrix, does not cause any chemical changes in the network structure and interacts between the bonds. It is observed that the % transmittance values decrease in peak intensities due to the increase in the reinforcement ratio. It can be said that possible interactions in the spectrum obtained with nano h-B_4C reinforcement at different rates to the two-component adhesive occur between aromatic groups in the polymer chains and that the conformational changes and peak stresses based on this shift to the low energy region. It indicates that the C = C tensions in the 1731-1616 cm^{-1} band observed as a mixture added with nano-reinforced, h-B_4C is distributed in the polymeric matrix. The results show that the h-B_4C particles settle in a certain proportion. Amorphous polymer matrix prepared with Araldite 2011 and MGS L285 show large peaks in between Figure 6 and Figure 8. Nano h-B_4C shows B-C and BC-B-C peaks at approximately 832 and 1552 cm^{-1} wavelengths, respectively [15,16]. Amine peaks are seen at 3381 cm^{-1} from Araldite 2011 B hardener [17]. The presence of aromatic C-H bonds originating from epoxy glue has been observed at wavelengths of 2965 and 2873 cm^{-1} [8,17,18]. It indicates that the C = C tensions in the 1608 cm^{-1} band range observed in the mixture with nano-reinforced mixture, h-B_4C is distributed in the polymeric matrix [8,17]. Hardener originated C-N bonds are seen at 695 cm^{-1} [15]. The presence of oxirane rings in the chemical formula of the epoxy at 3057, 915 and 835 cm^{-1} [8,17,19]. With h-B_4C nano reinforcement, in epoxy structure at 1363, 1190, 1063 and 1036 cm-1 wavelengths, respectively, C-CH3 (methyl origin), C-OH (secondary alcohol source), CO (ether sourced) and CO (aromatic ring sourced) bonds it was observed that peak vibrations changed [8,9,17,19,20].

According to the results obtained from FT-IR spectra, it has been shown that nano h-B_4C, which is specified as nano reinforcement in the polymer matrix, does not cause any chemical change in the network structure and interacts between bonds. It is observed that the % transmittance values decrease in peak intensities due to the increase in the reinforcement ratio in Figure 7 and Figure 8. This decreasing provides guidance that the amount of nano-reinforced materials incorporated into epoxy can be quantitatively determined using the FT-IR technique.

ACKNOWLEDGMENT

This study was supported by the Scientific and Technological Research Council of Turkey (TUBITAK) with 119M939 project number.

REFERENCES

[1] H Ulus, "Ark Deşarj Yöntemi ile Bor Nitrür Nano Levha Üretimi karakterizasyonu ve Karbon Kumaş/Epoksi Kompozitlerin Mekanik Özelliklerine Katkisinin İncelenmesi", *Master Thesis*, Selçuk University, Konya, Turkey, 2014

[2] W Y Mai, Z Z Yu, *Polymer Nanocomposites*, CRC Woodhead Publishing in Materials, Cambridge, 2006
[3] V K Thakur, M K Thakur, "Eco-Friendly Polymer Nanocomposites: Chemistry and Applications", *Advanced Structured Materials*, Springer: New Delhi, 2015
[4] Y Tsuji, Y Kitamura, M Someya, T Takano, M Yaginuma, M , K Nakanishi, K Yoshizawa "Adhesion of Epoxy Resin with Hexagonal Boron Nitride and Graphite", *ACS OMEGA*, vol 4,pp 4491-4504, 2019
[5] D Lee, S H Song, J Hwang, S H Jin, K H Park, B H Kim, S H Hong, S Jeon, "Enhanced Mechanical Properties of Epoxy Nanocomposites by Mixing Noncovalently Functionalized Boron Nitride Nanoflakes", *Small,* vol 9,pp 2602-2610, 2013
[6] K C Yung, H Liem, "Enhanced Thermal Conductivity of Boron Nitride Epoxy-Matrix Composite through Multi-Modal Particle Size Mixing", *Journal of Applied Polymer Science*,vol 106,pp 3587-3591, 2017
[7] Z Y Lin, A McNamara, Y Liu, K S Moon, C P Wong, "Exfoliated Hexagonal Boron Nitride-Based Polymer Nanocomposite with Enhanced Thermal Conductivity for Electronic Encapsulation", *Composite Science Technology,*vol 90,pp 123-128, 2014
[8] İ Akpinar, A Gürses, S Akpinar, K Gültekin, A Ozel, "Investigation of mechanical and thermal properties of nanostructure-doped bulk nanocomposite adhesives", *Journal of Adhesion,*vol 94,pp 847-866, 2018
[9] B M V Romão, M F Diniz, F P Margarete, V L Lourenço, L C Pardini, R C L Dutra, "Characterization of the Curing Agents Used in Epoxy Resins with TG / FT-IR Technique", *Polimeros,*vol 16,pp 94-98, 2006
[10] E Topcuoğlu, "Preparation and Characterization of Polymer Composites Containing Boron Compounds", Master's Thesis, Middle East Tchnical Universitey, Ankara, Turkey, 2016
[11] F Yuan, W Jiao, F Yang, W Liu, Z Xu, R Wang, "Surface modification and magnetic alignment of hexagonal boron nitride nanosheets for highly thermally conductive composites", *RSC Advances* pp 43380-43389,2017
[12] H Harrison, J T Lamb, K S Nowlin, A J Guenthner, K B Ghiassi, D Kelkar, R Alston, "Quantification of hexagonal boron nitride impurities in boron nitride nanotubes via FTIR spectroscopy", *Nanoscale Advances,* vol 2, pp 1693-1701,2019
[13] M Celiński, M Sankowska, "Explosibility, flammability and the thermal decomposition of Bisphenol A, the main component of epoxy resin", *Journal of Loss Prevention Process Industries,*vol 44,pp 125-131, 2016
[14] R Kumar, A Mishra, S Sahoo, B P Panda, S Mohanty, S K Nayak, "Epoxy-based composite adhesives: Effect of hybrid fillers on thermal conductivity, rheology, and lap shear strength", *Advanced Polymer Technologies,*vol 30,pp 1365-1374,2019
[15] A S Kipcak, P Gurses, E M Derun, N Tugrul, S Piskin, "Characterization of boron carbide particles and its shielding behavior against neutron radiation, *Energy Conversion Management*,vol 72,pp 39-44,2013
[16] J. Wang, Y. He, Z. Xie, C. Chen, Q. Yang, C. Zhang, B. Wang, Y. Zhan, T. Zhao, "Functionalized boron carbide for enhancement of anticorrosion performance of epoxy resin",*Polymer Advanced Technologies.*vol.29,pp.758-766, 2018.
[17] F. Contu, L. Fenzy, S.R. Taylor, "An FT-IR investigation of epoxy coatings as a function of electrolyte composition", *Progress in Organic Coatings,* vol.75,pp.92–96, 2012.
[18] P. Maity, S.V. Kasisomayajula, V. Parameswaran, S. Basu, N. Gupta, "Improvement in surface degradation properties of polymer composites due to pre-processed nanometric alumina fillers", *IEEE Transactions on Dielectrics and Electrical Insulation,* vol.15, pp.63-72,2008.
[19] T. Yalçınkaya, "Synthesis and Characterization of Epoxy / Boron Nitride Composite for Aerospace Applications", *Journal of Aeronautics and Space Technologies,*vol.12,pp.87-94, 2019.
[20] J Abenojar, M A Martínez, F Velasco, "Effect of Boron Carbide Filler on the Curing and Mechanical Properties of an Epoxy Resin", *The Journal of Adhesion,* vol 85,pp 216-238, 2009

Technical and Economical Assessment of Liquid Petroleum Gas Modification/Conversion on a Gasoline Engine in Turkey for Time Interval between 2016-2020

E.CANLI[1]

[1] Selcuk University, Konya/Turkey, ecanli@selcuk.edu.tr

Abstract **– A 1.2-liter gasoline spark ignition internal combustion engine of a Renault Symbol sedan vehicle was adapted to Liquid Petroleum Gas (LPG) fuel system in 2016. The LPG system utilize LPG and gasoline fuels according to the operational needs. After modification, fuel consumption of the engine, fuel prices per liter, travelled distance in kilometers, dates and some additional data were recorded till November 2020. Collected data is processed for derivation of performance and economy indicators. Literature and web have been reviewed in order to find related works. Providing a solid application example of LPG conversion is aimed in this work. Also, a snapshot of fuel prices between 2016 and 2020 is being presented by graphical resolution. Approximately 12 to 15 times of initial investment cost has been returned till November 2020 by the LPG system. Also, slightly less CO_2, i.e. 6%, has been emitted by the utilization of the system. Time resolution of the data is provided by graphics and integral data is also given. LPG conversion is found beneficial in terms of fuel cost. Also, viewed four years period does not reveal any performance problems in terms of energy spending per kilometer. However, engine additional amortization due to LPG system is not considered and this should be kept in mind.**

Keywords **– Combustion, consumption, conversion, emission, fuel cost, gasoline, LPG, internal combustion engine.**

I. Introduction

THE fuel consumption has been a hot topic for nearly half century after fuel crisis in 70s. Prices and environmental impacts lead to numerous works in order to reduce fuel consumption or find alternatives. LPG utilization in gasoline engines not only has operational cost advantage and slight environmental advantages but also find a solution for a byproduct of petroleum processing. Some related works in the context of the present paper are summarized below.

A 1992 Chevrolet Lumina was tested with three different LPG conversion systems for emissions and economy according to US Federal Test Procedure [1]. Authors tested four different fuel blends with the three different LPG kits. They reported variability between emission values, though emission amounts and compositions were found proper. In a similar work, authors used Renault brand gasoline spark ignition engine in order to test LPG systems in terms of torque and exhaust emissions [2]. They reported a decrease of 3% for torque while exhaust emissions were improved. However, they studied on engine mapping for best performance. An 80000 km endurance test were stated to be planned for their future work. In another work, authors tried to analyze the difficulty of presenting a universal LPG system that regulate all available vehicles in the market [3]. They compared laboratory test systems and a simulation with commercially available conversion kit. The investigated kit uses exhaust oxygen sensor in order to give feedback to LPG system. Emissions were studied before and after catalyst convertors. Authors suggested that additional measures were required for different vehicles while it was very hard to provide a single kit for all vehicles. Transferring LPG to combustion chamber is also investigated since carburetion and injection lead to different performances [4]. Authors suggested that injecting LPG is superior comparing to carburetion. LPG blends with biogas and gasoline is investigated in order to achieve better and favorable exhaust emissions [5]. However, authors report serious decrease in brake thermal efficiency and increase in brake specific fuel consumption. Various market numbers worldwide also show a tendency for LPG, strengthening its current position as an alternative [6]. Gumus reported four different ratios of LPG in spark ignition internal combustion gasoline engine in order determine engine performance in respect of exhaust emissions, fuel consumption and volumetric efficiency [7]. Author states that emissions are favorable with LPG utilization. However, volumetric efficiency decreases due to LPG. On the other hand, 25% LPG maintains thermal efficiency with a lower air to fuel ratio. Erkus et al. investigated ignition timings to show changes in heavy load performance of LPG in gasoline engine [8]. Authors suggest advancing ignition timing to achieve better performance. However, advancing ignition timing also increases emission, according to the authors. Mustaffa et al. investigated the effects of injecting LPG into combustion chamber in liquid phase and compared results with gasoline [9]. They compare peak pressure values in combustion chamber. As a result, they reported better combustion stability with liquid injection of LPG comparing to gasoline. Liquid injection of LPG is not compared with gaseous injection of LPG in that work. According to the literature survey in the present work, it is

seen that most research are based on controlled laboratory testing. Therefore, LPG usage data from application and field for relatively long period of time will be beneficial. This work provides such data.

In this work, data from 4 years period for LPG usage in gasoline spark ignition internal combustion engine of a sedan car is presented. The data consist of fuel prices, travel kilometers, fuel purchase dates and derived magnitudes. Following section gives information about the vehicle and the engine. Method for data collection and derived magnitudes are provided. In the results part, data depending on date are presented by graphics. Derived magnitudes are listed. Results are evaluated in terms of investment and engine performance. Some conclusions and remarks are given in the last part.

II. Technical Information and Method

Present data consist of fuel consumption of 2013 Renault SR Symbol Joy 1.2 lt 16 valve 75 hp D4F engine type vehicle between 01.11.2016 and 30.10.2020. Fuel stations were in Turkey, mostly in Konya. Engine actual volume is 1.149 lt. Vehicle type is M1 in Turkish standards. The aerodynamic form is typical to sedan vehicles. Net weight of the vehicle is 1114 kg and maximum allowable gross weight is 1505 kg. Although engine nominal power is 75 hp, legal permit engine power is 55 kW.

Figure 1: An illustrative example of the vehicle type and brand.

The vehicle was used without any fuel system modifications between 02 July 2013 and 01 November 2016. The average gasoline fuel consumption per kilometer was measured as 0.065 lt/km in that time interval. The reason for keeping engine without LPG conversion was to protect engine warranty which was three years from the purchase of the car. The engine fuel system of the vehicle was modified in order to add a Liquid Petroleum Gas (LPG) fuel system in November 2016. ATIKER AUTOGAS SYSTEMS & LPG TANKS was selected as the system manufacturer. The company origin is Turkey. LPG kit model was MicroFast. SR07 model regulator took place in the kit. 40 lt LPG storage was selected. Vehicle total km usage was 62305 km at the time of modification. This LPG system injects LPG into intake manifold by injectors. There are four injectors for four cylinders. The LPG system is very robust in terms of maintenance. Only LPG filter is needed to be changed at every 10000 km usage. Also, parameters in electronic control unit of LPG system is monitored during maintenance. Maintenance service can change some parameter values accordingly.

Initial cost of the LPG system was 1,100.00 TL including labor cost. Since the initial implementation, no operation anomaly or symptom in engine has been observed. Total operation time is 1439 days and 86888 km.

Fuel consumption data were comprised by the following algorithm:

1. Gasoline and LPG fuel tanks are filled completely for the first time just after LPG conversion.
2. Travel kilometer indicator is reset (set to zero).
3. Next fuel purchase is done in the shortest possible time, just after the LPG fuel indicator shows the minimum fuel.
    a. LPG and gasoline tanks are filled full.
    b. Fuel filling stops when fuel pumps automatically stop.
    c. Travel kilometer, date of the fuel purchase, fuel prices per liters are recorded.
    d. Travel kilometer indicator is reset (set to zero).
4. Purchase receipt is numerically, physically and visually stored.
5. Data occur in step 3 is transferred to a calculator.
6. The amount of fuel purchase in step 3 is related to the travel kilometer that is recorded in step 3.
7. Fuel costs per kilometer, fuel consumption per kilometer, total purchase prices, price that would occur if LPG was not adapted, price differences between fuels, savings based on LPG system investment are calculated.
8. Return to step 3.

Step 3-point b is emphasized because this is a good way of ensuring a limit for fuel tanks to be full at same level with their previous full levels. LPG pumps have a pressure limit for stopping. Gasoline pumps stop when fuel level reach to a height that is almost same each time. Step3-point b is so important that if there would be, let's say 10 percent uncertainty in this level, whole derivatives would have this uncertainty.

The amount of fuel purchase in step 3 is related to the travel kilometer that is recorded in step 3 since fuel levels are returned to their previous level where travel kilometer was set to zero. Therefore, the purchased amount of fuels corresponds to the travel kilometer.

The algorithm needs fuel filling to the storages till filling limit is reached each time of fuel purchases. Nevertheless, if total sampling period is long as in this work, amount of data compensates fuel purchase problems relating to mentioned rule. For instance, gasoline purchase per LPG purchase is about 2 to 3 liters. Therefore, purchasing gasoline each time LPG purchased can be burdensome. Again, the data would compensate skipped fuel purchases if consequent gasoline purchase would be done fulfilling limit.

All receipts were recorded. It may be beneficial to mention that there are several mobile phone applications for this task. However, author of the present work used MS-Excel in order to store data and make calculations.

Calculations for economical evaluations are based on data of Turkish Republic Central Bank (Türkiye Cumhuriyeti Merkez Bankası) [10]. Daily currency rates, monthly gold prices and yearly interest rates are taken from the bank data. United States Dollar (USD) is selected for the comparison currency against Turkish Lira (TL). Republic Gold (Cumhuriyet Altını) price is used for gold price. Interest rates are yearly deposit account interest rates. Pay back periods are calculated according to changes in currencies, gold prices and interest rates.

Average weight of 1309.5 kg (average of net and maximum allowable gross weights) is used for scaling energy consumption results. This is important since machines cannot be compared based on their energy consumption without considering their work outputs. In other words, energy efficiency is the proper comparison tool. However, determining the work done by the present vehicle is not possible with present data, though travel kilometer, average weight and energy consumption are known. doing travel work during calculations.

Energy consumption is calculated using lower calorific values of fuels. Lower calorific values of the fuels were taken from a web source [11]. It is known that fuel contents and densities may change according to date and region. Therefore 10% uncertainty exists in derivative results.

Chemical formulas for fuels are assumed as following:

$$C_3H_8 \rightarrow \text{LPG} \quad (1)$$

$$C_8H_{18} \rightarrow \text{Gasoline} \quad (2)$$

$$CH_4 \rightarrow \text{NaturalGas} \quad (3)$$

Carbon dioxide emission due to the energy conversion by combustion in engine is approximated by stochiometric combustion of assumed fuel chemical formulas.

Calculating savings due to LPG conversion investment is done by considering 0.065 lt/km average gasoline usage, which is based on measurement before the LPG modification.

The LPG system has an operational life of ten years according to product catalogue. In the present stage, additional amortization cost of the engine due to LPG conversion is absent. Reader should be aware of this fact.

III. Results

Modified vehicle for LPG fuel system has been operating for 1439 days and 86888 km. 178 sampling (fuel purchases) have been done for the time interval, which is between 01.11.2016 – 30.10.2020. Considering average weight and operation kilometer, following results are derived from the dataset:

- Total energy consumption is 164042.5 MJ based on lower calorific values of the fuels. 1.44 kJ/kg·km is realized by the modified vehicle.
- 5388.65 lt LPG, 1012.29 lt gasoline were consumed. Total liquid fuel consumption is 6400.94 lt. 5.62×10^{-5} lt/kg·km is realized by the modified vehicle.
- 2367.57 kg CO_2 was emitted due to LPG fuel.
- 628.26 kg CO_2 was emitted due to gasoline fuel.
- Total CO_2 emission is 2995.83 kg. 2.63×10^{-5} kg/kg·km is realized by the modified vehicle.
- 2949.81 kg CO_2 would be emitted if only LPG was used. Carbon to total weight ratio is assumed 0.81.
- 3182.97 kg CO_2 would be emitted if only gasoline was used. Carbon to total weight ratio is assumed 0.84.
- 2612.14 kg CO_2 would be emitted if only natural gas was used. Carbon to total weight ratio is assumed 0.75.
- Fuel liter per kilometer values are: 0.0737 lt/km for gasoline+LPG; 0.062 lt/km for LPG; and 0.0117 lt/km for gasoline.
- Fuel cost per kilometer values are: 0.247 TL/km for gasoline+LPG; 0.177 TL/km for LPG; and 0.069 TL/km for gasoline.
- Total fuel cost of the system is 21,451.00 TL (15,433.00 TL for LPG and 6,017.00 TL for gasoline).
- Total money savings due to LPG system is 12,756.00 TL or 11.59 times the cost of the LPG system. Net Present Value (NPV) of the total savings is 14,116.00 TL based on recent prices; and 17,195.00 TL based on yearly interest rates.
- Saving per kilometer by the LPG system is 0.146 TL/km based on actual spending; 0.162 TL/km based on recent fuel prices; and 0.197 TL/km based on yearly interest rates.
- Initial cost payback periods are 121 days based on actual payments; 128 days based on interest rates; 132 days based on currency investment and 135 days based on gold investment.

Figures 2 to 6 are given in order to provide a time resolution of the collected data. Figure 2 shows fuel consumption in liter per traveled kilometer. Red marks show dates of gasoline purchases. Blue marks show LPG purchases. When red marks are absent while blue ones are present, it means that gasoline purchase is skipped. Those instances are compensated by gasoline purchases where green marks (total liter per kilometer) show a higher value comparing to the average. Similarly, occasional markers below and above average trends are due to skipped fuel purchases. Nevertheless, the big portion of numerous samples constitute o clear trend. There is a slight decrease in fuel consumption as time passes. Although the reason cannot be determined without additional work, it is

thought that LPG utilization may lead to changes to some extend in mechanical properties of engine parts, especially combustion chamber.

Figure 3 shows that price gap between gasoline and LPG is increasing by time. Therefore, savings due to LPG usage is increasing. There is a linear proportionality as expected.

Figure 4 shows a slight increase in travel cost per kilometer although Figure 2 shows a slight decrease in fuel consumption per kilometer. This is due to fuel price increase in TL as shown in Figure 5 and 6. Gasoline and LPG fuel prices have been decreasing starting from early parts of 2018 in USD. The devaluation of TL leads to slight travel cost increase in Figure 4.

IV. Conclusion

This work exhibits a particular LPG utilization experience in a personal gasoline vehicle by means of collected data during approximately 1500 days interval. Considering the approach given in second section; following conclusions are made.

- LPG modification of the gasoline engine is a good investment that paybacks in one third of one year. This period is 3% of the operational life of the LPG system. However, additional amortization cost of the engine due to LPG conversion is not known.
- LPG system also offers slightly less carbon dioxide emissions about 6% comparing to gasoline engine.
- Fuel consumption of the vehicle decreased slightly which can be attributed to a possible change in material properties of the engine. Especially, surfaces of combustion chamber is thought to be effected.
- Price difference between LPG and gasoline seems increasing which leading to increasing savings.
- Although TL/km fuel cost seems increasing, slightly decreasing fuel consumption and decreasing fuel prices based on USD shows a potential for LPG.

V. Disclosure

All data, derived results and evaluations are based on free, individual and self-experiences. Data and results are subjective and cannot be generalized. None of the content is at a state of binding for any individuals or institutions of any kind. In the present stage, additional amortization cost of the engine due to LPG conversion is absent. Reader should be aware of this fact.

References

[1] E Bass, B Bailey, and S Jaeger, "LPG Conversion and HC Emissions Speciation of a Light-Duty Vehicle," *SAE Technical Paper 932745*, 1993, https://doi org/10 4271/932745

[2] A Gerini, G Monnier, and R Bonetto, "Ultra Low Emissions Vehicle Using LPG Engine Fuel," *SAE Technical Paper 961079*, 1996, https://doi org/10 4271/961079

[3] W Smith, D Timoney, and D Lynch, "Emissions and Efficiency Comparison of Gasoline and LPG Fuels in a 1 4 Litre Passenger Car Engine," *SAE Technical Paper 972970*, 1997, https://doi org/10 4271/972970

[4] B Erkus, A Sürmen, and M I Karamangil, "A comparative study of carburation and injection fuel supply methods in an LPG-fuelled SI engine", *Fuel,* vol 107, pp 511–517, 2013

[5] S Simseka, and S Uslu, "Investigation of the impacts of gasoline, biogas and LPG fuels on engine performance and exhaust emissions in different throttle positions on SI engine", *Fuel*, vol 279, 118528, 2020

[6] R Biscoff, M Akple, R Turkson, and W Klomegah, "Scenario of the emerging shift from gasoline to LPG fuelled cars in Ghana: A case study in Ho Municipality, Volta Region", *Energy Policy*, vol 44, pp 354–361, 2012

[7] M Gumus, "Effects of volumetric efficiency on the performance and emissions characteristics of a dual fueled (gasoline and LPG) spark ignition engine", *Fuel Processing Technology*, vol 92, pp 1862–1867, 2011

[8] B Erkus¸ M I Karamangil, and A Sürmen, "Enhancing the heavy load performance of a gasoline engine converted for LPG use by modifying the ignition timings", *Applied Thermal Engineering*, vol 85, pp 188-194, 2015

[9] N Mustaffa, M Fawzi, A Khalid, S A Osman, N Jaat, and M M Tukiman, "Effects of Liquid LPG Injection on Combustion Stability in Spark Ignition Engine", *Journal of Advanced Research in Fluid Mechanics and Thermal Sciences*, vol 76(1), pp 164-171, 2020

[10] https://evds2 tcmb gov tr/index php?/evds/serieMarket/#collapse_2 [Access time: 03 11 2020]

[11] https://www engineeringtoolbox com/fuels-higher-calorific-values-d_169 html [Access time: 03 11 2020]

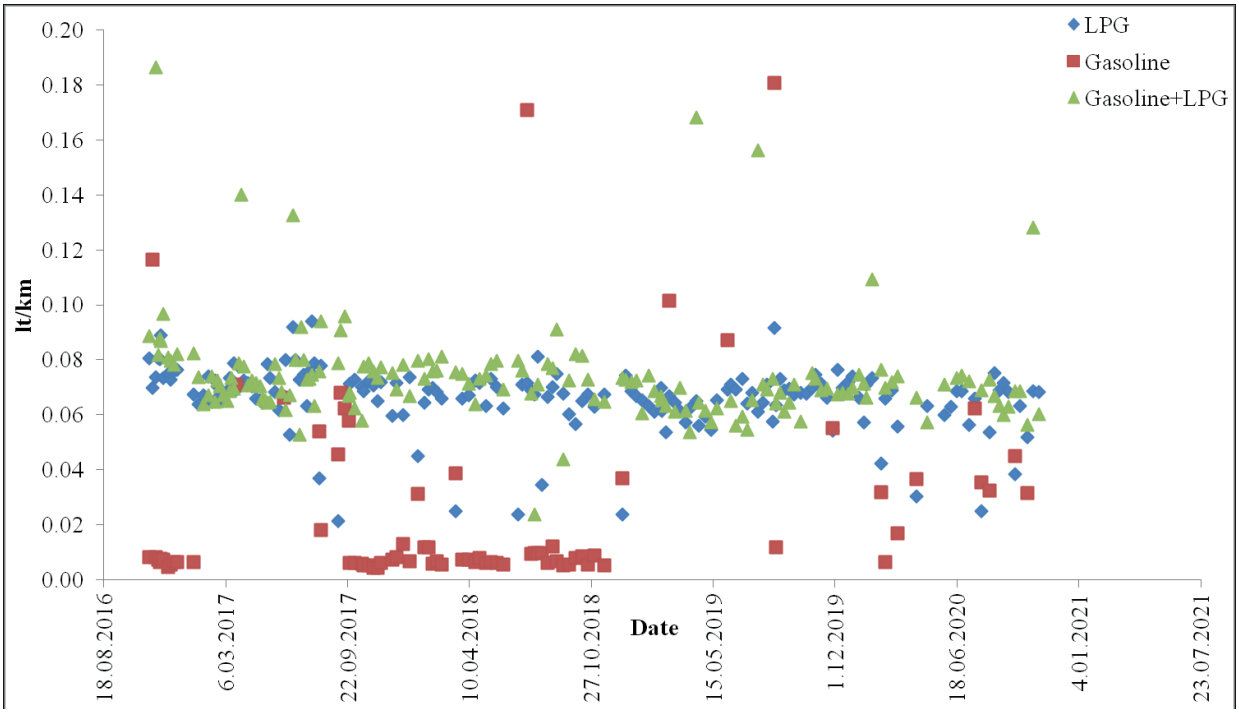


Figure 2: Liter per kilometer fuel consumption.

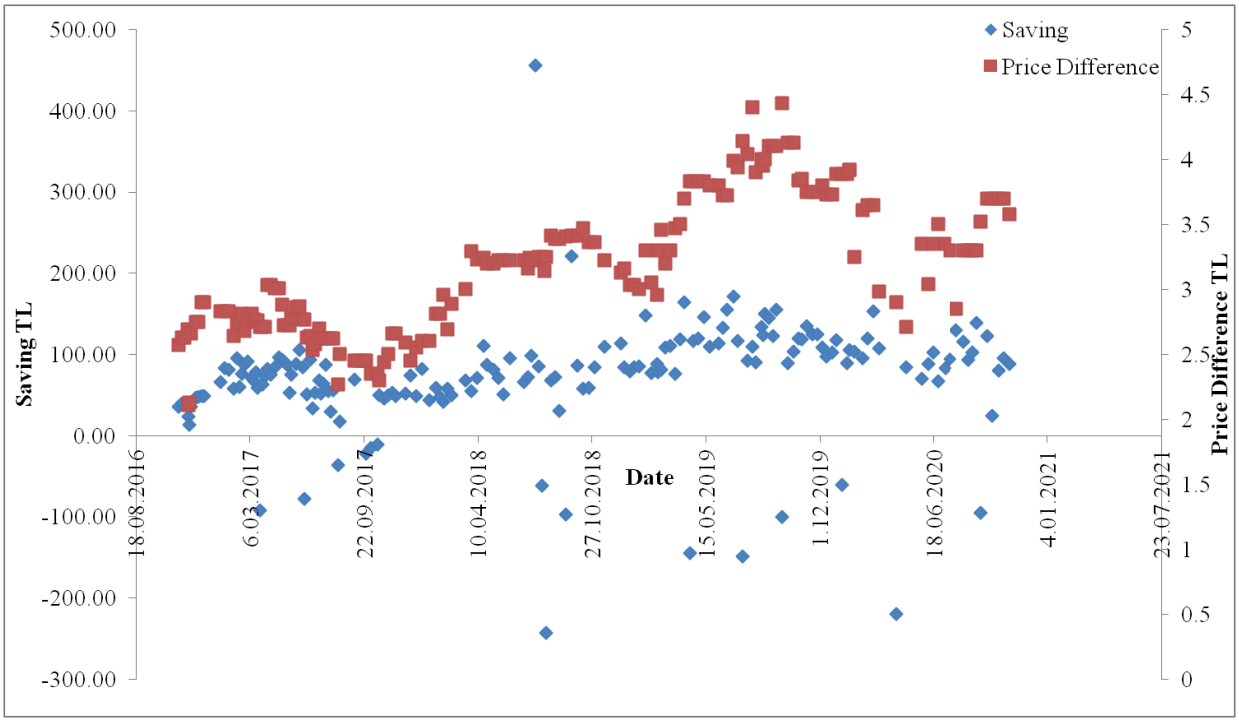


Figure 3: Price difference between LPG and gasoline; and savings per period.

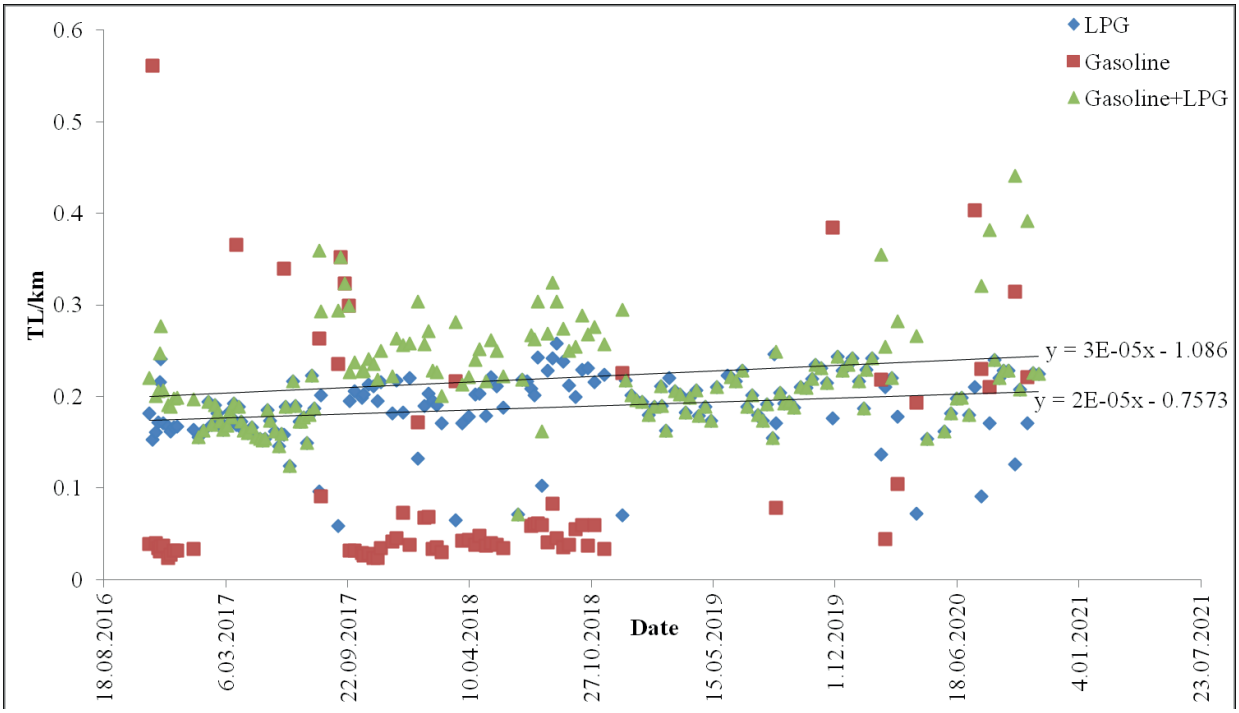


Figure 4: Fuel cost per kilometer.

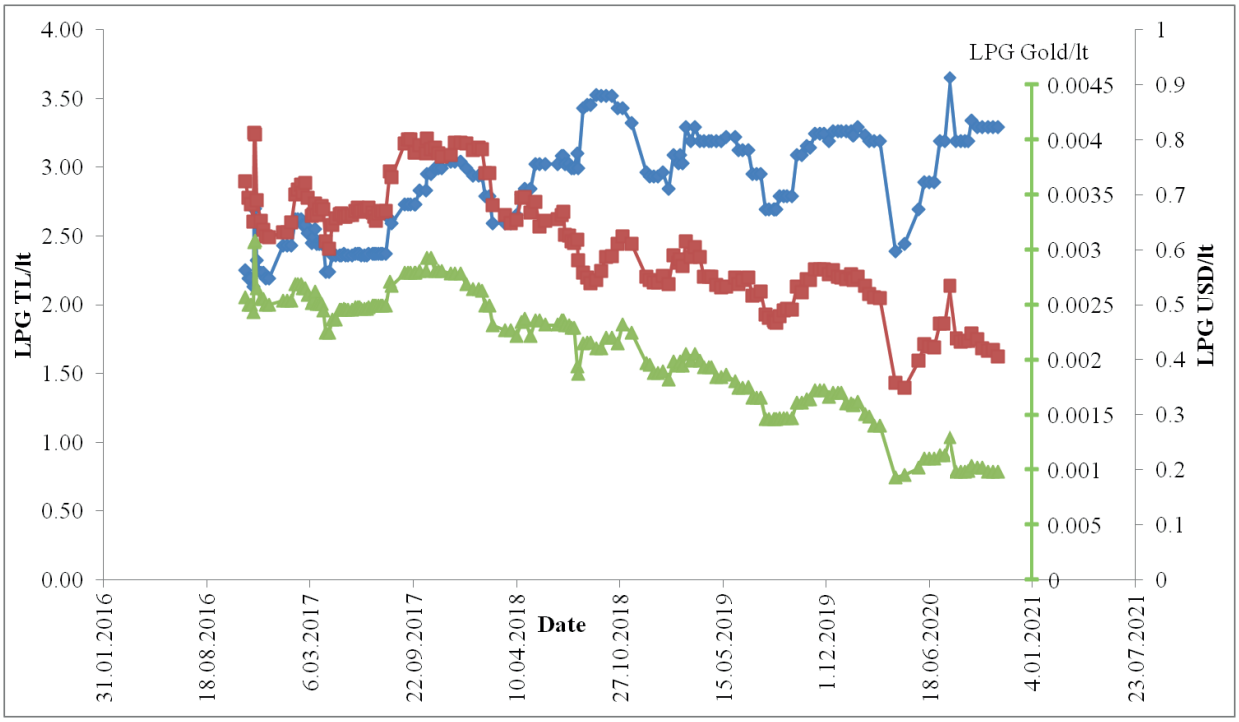


Figure 5: LPG cost per liter.

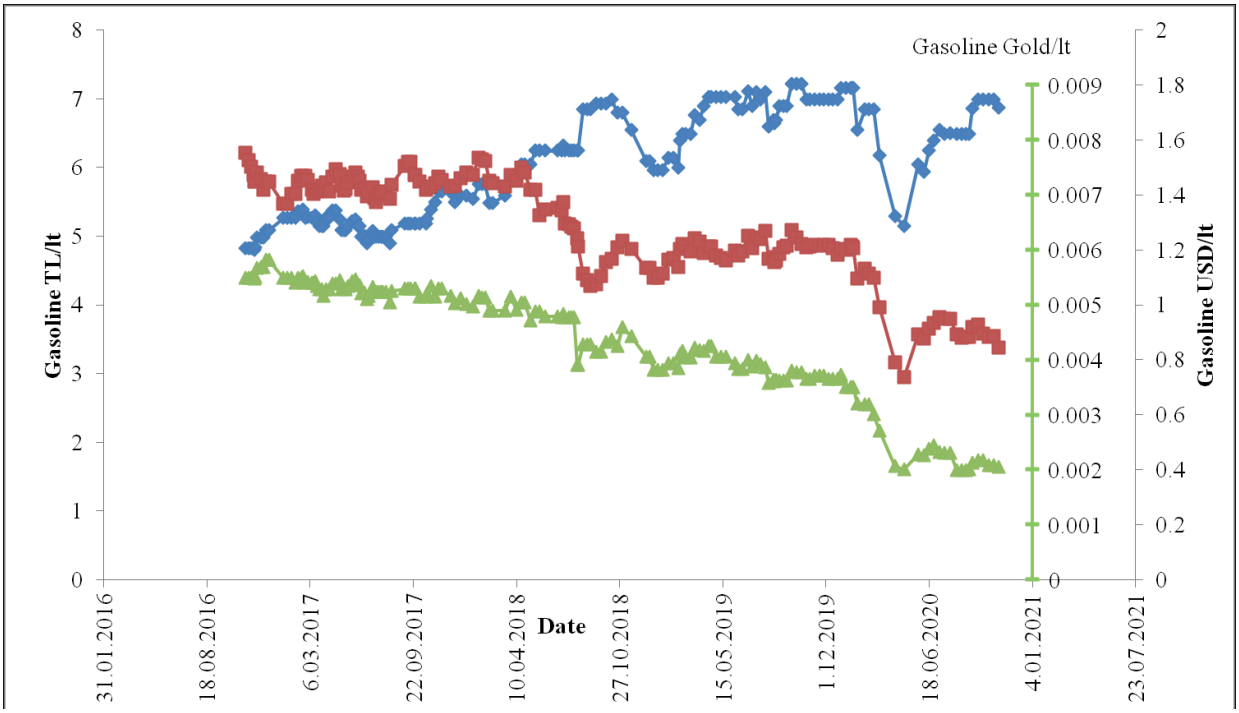


Figure 6: Gasoline cost per liter.

An Investigation of Factors Affecting Machinability of Milling Toolox 44 Hot Work Tool Steel by Taguchi Method

R. BİNALİ[1], H. DEMİR[2], S. NEŞELİ[1], S. YALDIZ[1]

[1] Selcuk University, Konya/Turkey, rustem.binali@selcuk.edu.tr
[2] Karabük University, Karabük/Turkey, hdemir@karabuk.edu.tr
[1] Selcuk University, Konya/Turkey, sneseli@selcuk.edu.tr
[1] Selcuk ,

Abstract **- In the study, the machinability of the Toolox 44 (44 HRc) hot work tool steel material, which is used in military land vehicles and produced as pre-hardened (with four different feed rates, four different cutting speeds and two different depths of cut parameters) was investigated using milling method. According to the Taguchi $L_{32}(4^2x2^1)$ experimental design, physical tests were carried out in which response values of cutting force and surface roughness were obtained. In the chip removal process, a single type TiAlNi coated carbide cutting tool was used, and the cutting forces were measured with a Kistler 9257B type dynamometer and the surface roughness was measured with a Mahr (Marsurf PS1) table type roughness device. Statistical analysis was made depending on the "lowest and best" objective function through the Minitab 18 package program. It was concluded that the optimum processing parameters obtained as a result of the experimental study were 180 m/min cutting speed, 0.4 mm chip depth of cut and 0.4 mm/tooth feed rate. In addition, the lowest surface roughness value was obtained as 0.533 µm and the highest surface roughness value was obtained as 3,126 µm. In cutting forces, the highest value was obtained as 410.7 N and the lowest cutting force value as 16.3 N. In addition, according to the results of ANOVA analysis applied, it was observed that the most effective factor on Fz was the amount of feed rate, while the most effective factor on Ra was the cutting speed.**

Keywords **- TOOLOX 44, Milling, TAGUCHI Method, Surface Roughness, Cutting Force.**

I. INTRODUCTION

Machining is one of the most common methods used to make raw materials into the desired end product, give the final shape, and remove undesirable places in the raw material. The hardness of the cutting tool used in machining to reach the final product should be higher than the cut material. Turning, milling, threading and drilling etc. used in industry machining principles are taken into account in all of the operations.

Nowadays, especially milling operations are widely preferred in the industry due to their high axial movement ability. The cutting tool used in milling has more than one cutting edge and cutting edge, and each has a chip removal capacity from the material. With this method, it is easier and more advantageous to make the chip removal process from the material more efficient, to have a smaller surface roughness value, to perform more precise processes and to produce products within tolerances [1].

Face milling, which is one of the milling processes, has a wide area of use in the machining sector, in the production of machinery and parts, in the aviation, automotive and defense industries. For this reason, the face milling method is widely used [1,2]. Since the quality of the surface that occurs as a result of the face milling process is a factor that significantly affects the production costs and the quality of the product, the study was also examined through this method [3].

The most important factors in performing machining are processing parameters. These parameters are cutting speed, feed rate and depth of cut. Every change in these parameters causes changes in the cutting force. Cutting force increases as the feed and depth of cut increase. As the cutting speed increases, the cutting force decreases. For this reason, these parameters should be selected well during the milling process [4]. These parameters are also the most important parameters that determine the machinability of the material. Other factors affecting machinability are the vibration of the machine, the hardness and ductility of the workpiece material, the geometry of the cutting tool material and the cutting conditions [5].

Predictive studies are carried out to obtain appropriate parameter combinations and to determine the interactions between processing parameters [6]. Considering the industrial conditions, it is an efficient and important approach to apply estimation technique and experimental design methods in order to ensure that the experimental working time is limited and the economic conditions can be done in the most efficient way and the results obtained can be interpreted correctly and the relationships between parameters can be determined. In addition, all other methods applied in order to increase productivity and quality have a guiding and supportive role. Important considerations in experimental estimation design technique are reliability, cost, and sensitivity [7-9]. Taguchi method is the statistical method used to obtain the best experimental results with a minimum number of experiments. Thus, cost and time savings are achieved by reducing the number of experiments [3,10,11]. Because of these advantages, it is the most widely used experimental design method by researchers to determine the relationships between different levels of different parameters [6,12].

In this study, the effects of depth of cut, cutting speed and feed rate parameters on cutting force and surface roughness were evaluated using Taguchi method in milling Toolox 44 hot work tool steel. In the experiments, Taguchi L_{32} orthogonal array was used, and the optimization studies determined the most suitable processing parameters giving the lowest cutting force and surface roughness of the control factors. In addition, the interaction of all control factors with each other was determined by signal to noise ratio (S/N) and analysis of variance (ANOVA).

II. MATERIALS AND METHOD

Machinability experiments of the study were carried out using the Johnford VMC-550 vertical machining center bench. Toolox 44 steel used in experimental studies is used in the production of plastic injection molds, forming and cutting molds, metal injection molds, hot tool steels and hot forging molds. The chemical composition of Toolox 44 material is given in Table 1 and its mechanical properties are given in Table 2.

Table 1: Chemical composition [13].

Material	C	Si	Mn	P	
Toolox44	0,30	0,61	0,89	0,010	
	S	Cr	Ni	Mo	V
	0,0009	1,23	0,66	0,79	0,145

Table 2: Mechanical properties [13].

Tensile Strength R_m N/mm^2	Yield Strength Rp 0,2 N/mm^2	Elongation %	Hardness HBW
min. 1380	min. 1200	min. 10	min. 410

As the cutting tool used during the experimental studies, TiAlN coated carbide cutting tools using the KENNAMETAL WPGX 030204 LD 080 physical coating method and the tool holder coded 16Y02R060A16SWP03L150 were used. Measurement of surface roughness (Ra) was measured with Marsurf PS1 brand table surface roughness device, which uses tracer tip instruments method. After each experiment, surface roughness measurements were made from 3 different points of the processed surface and their average values were used for evaluation. Cutting forces were measured by using a Kistler 9257B type dynamometer.

Experiments were carried out under 2 different depths of cut, 4 different cutting speeds and feed rates and dry cutting conditions. Processing parameters in the experiments are given in Table 3.

Table 3: Machining parameters.

Material	Deep of cut (mm)	Feed rate (mm/tooth)	Cutting speed (mm/dk)
Toolox 44	0,2-0,4	0.4-0.8-1.2-1.6	150-180-210-240

The experimental design was carried out using the Taguchi L_{32} orthogonal array with the processing parameters in Table 3. Taguchi L_{32} experiment design (first 3 columns) and physical test results (4th and 5th columns) prepared with Minitab 18 software are given in Table 4.

III. TAGUCHI METHOD

The main factor that makes up the difficulty of experimental studies is the experimental designs that contain combinations of processing parameters. Experimental sequences that can occur depending on the number of parameters can produce extremely complex and time-consuming results. This increases manufacturing costs and decreases productivity [14].

Taguchi method; It is widely used in engineering analysis due to the reduction in the number of experiments and lowering production and test costs. However, it offers a simple, efficient and systematic approach to determining optimum processing conditions [15,16]. The criterion used to measure and evaluate quality characteristics in the Taguchi experimental design method is the S/N ratio. S (Signal) indicates the actual value given by the system and desired to be measured, and N (Noise) indicates the share of uncontrollable unwanted factors in the measured value. There are 3 types of S/N ratios [17]. These are given in equations 1, 2 and 3.

The nominal-the best: $S/N=10\log(\bar{y}/S_y^2)$ (1)

The larger-the better: $S/N=-10\log(1/n\sum_{i=1}^{n}\frac{1}{y_i^2})$ (2)

The smaller-the better: $S/N=-\log(1/n\sum_{i=1}^{n}y_i^2)$ (3)

In the equations; $\bar{y}$ the average of the observed data, S_y^2 the variance of y, n the number of observation experiments and y the observed data [18].

IV. EXPERIMENTAL RESULTS AND DISCUSSION

The cutting force values and average surface roughness values obtained as a result of the experimental designs and experimental studies are given in Table 4.

Table 4: Taguchi L_{32} orthogonal experiment design and experimental results.

Deep of Cut (mm)	Feed Rate (mm/tooth)	Cutting Speed (mm/dk)	Surface Roughness Ra (μm)	Cutting Force (N)
0,2	0,4	150	0,946	198,3
0,2	0,4	180	0,533	143,3
0,2	0,4	210	0,952	103,9
0,2	0,4	240	2,127	16,3
0,2	0,8	150	1,691	218,5
0,2	0,8	180	1,184	153,3
0,2	0,8	210	1,454	117,3
0,2	0,8	240	1,616	58,2
0,2	1,2	150	2,264	294,4

0,2	1,2	180	3,126	255,2
0,2	1,2	210	1,013	201,3
0,2	1,2	240	1,876	175,6
0,2	1,6	150	2,448	292,4
0,2	1,6	180	2,806	339,4
0,2	1,6	210	2,585	303,6
0,2	1,6	240	2,067	271,3
0,4	0,4	150	0,583	212,6
0,4	0,4	180	0,612	198,7
0,4	0,4	210	0,677	178,1
0,4	0,4	240	0,551	150,8
0,4	0,8	150	2,236	316,5
0,4	0,8	180	2,216	279,1
0,4	0,8	210	2,280	263,9
0,4	0,8	240	2,175	247,9
0,4	1,2	150	2,010	359,5
0,4	1,2	180	1,937	326,8
0,4	1,2	210	1,556	340,6
0,4	1,2	240	1,991	293,5
0,4	1,6	150	2,655	370,8
0,4	1,6	180	2,768	410,7
0,4	1,6	210	2,565	368,0
0,4	1,6	240	2,094	316,6

In order to determine the optimum combinations of processing parameters, an experimental design was carried out orthogonally in order to share the effect levels of each parameter. In the study, in order to find the optimum points of the processing parameters, the S/N ratios were determined by considering the smaller-the better quality characteristic objective function (Table 5 and Table 6) [19].

Table 5: Average S/N response for Ra.

Level	Depp of Cut (mm)	Feed Rate (mm/tooth)	Cutting Speed (mm/min)
1	-4,219	**2,167**	-4,475
2	**-3,990**	-5,157	-4,012
3	-	-5,525	**-3,372**
4	-	-7,902	-4,559
Delta	0,228	10,070	1,186
Rank	3	1	2

Table 6: S/N response for cutting force.

Level	Depp of Cut (mm)	Feed Rate (mm/toot)	Cutting Speed (mm/min)
1	**-44,22**	**-41,84**	-48,81
2	-48,92	-45,30	-47,88
3	-	-48,74	-46,61
4	-	-50,40	**-42,99**
Delta	4,70	8,56	5,82
Rank	3	1	2

The parameter with the highest Delta values given for each level in Table 5 and Table 6 has the greatest effect on surface roughness and cutting force. The effect rankings are expressed as Rank in the table. Accordingly, in order to reach the optimum Ra level, it was determined that the possible parameter combination is 2nd level depth of cut, 1st level feed amount and 3rd level cutting speed. Likewise, it was determined that the ideal parameter combination for the cutting force is the 1st level cutting depth, 1st level feed and 4th level cutting speed.

In Figure 1 and Figure 2, graphs of S/N ratio responses for Ra and cutting force are given. The places where the slopes found on the graphs are the highest show the effectiveness levels of the parameters.

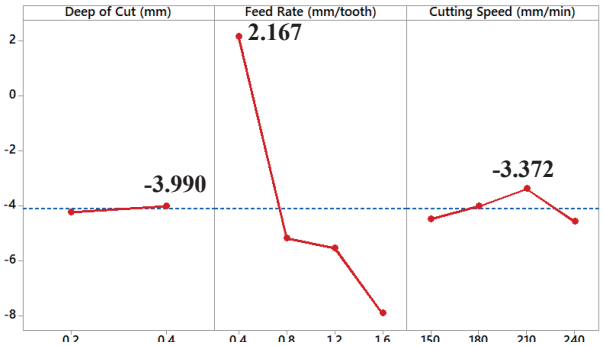


Figure 1. S/N chart for Ra.

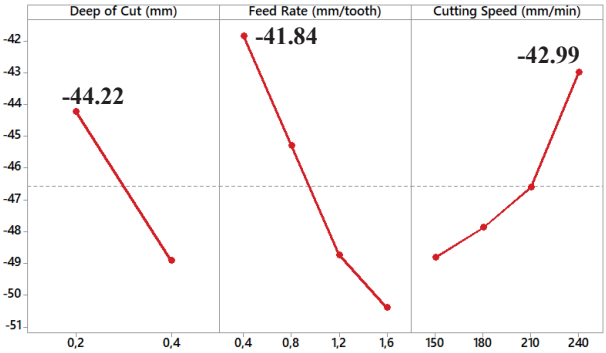


Figure 2. S/N plot for cutting force.

In the study, the ANOVA analysis method was used to determine and express the interactions of the processing parameters and control factors with each other more clearly and numerically (Table 7 and Table 8) [18].

Table 7. ANOVA analysis for Ra.

	Df	SS	MS	F	p	PCR (%)
Deep of Cut (mm)	1	0 417	0 417	0 07	0 791	%0 000 6
Feed Rate (mm/tooth)	3	455 07 2	151 69 1	27 11	0 000	%68,70
Cutting Speed (mm/min)	3	7 108	2 369	0 42	0 741	%1 07
Deep of Cut (mm)*Feed Rate (mm/tooth)	3	64 665	21 555	3 85	0 050	%9 76
Deep of Cut (mm)*Cutting Speed (mm/min)	3	12 571	4 190	0 75	0 550	%1 89
Feed Rate (mm/tooth)*Cuttin g Speed (mm/min)	9	72 156	8 017	1 43	0 300	%10 89
Error	9	50 362	5 596			
Total	31	662 35 3				100
R^2=%92 40						

Table 8: ANOVA analysis for cutting force.

	Df	SS	MS	F	p	PCR (%)
Deep of Cut (mm)	1	176 47	176 468	22 99	0 001	%18 6 5
Feed Rate (mm/tooth)	3	346 68	115 560	15 05	0 001	%36,6 4
Cutting Speed (mm/min)	3	156 42	52 142	6 79	0 011	%16,5 3
Deep of Cut (mm)*Feed Rate (mm/tooth)	3	42 59	14 197	1 85	0 209	%4 5
Deep of Cut (mm)* Cutting Speed (mm/min)	3	66 47	22 156	2 89	0 095	%7 02
Feed Rate (mm/tooth) *Cutting Speed (mm/min)	9	88 37	9 819	1 28	0 360	%9 34
Error	9	69 09	7 677			
Total	31	946 10				
R^2=%92 70						

In Table 7 and Table 8, p values expressing the significance level of the results of each variable, degree of freedom DF, sum of squares SS, mean of squares MS, F values and PCR percentage effect rates are given. In ANOVA, it is decided whether each parameter is effective on each other or not by looking at the p value. Considering the 95% confidence interval, it is concluded that the parameter is effective on the response when $p<0.05$ (5% significance value). The effect of control factors is determined by comparing the F values. Whichever factor has the highest F value, it is the factor that affects the result the most [2]. In this case, as a result of the examination of Table 7, it was concluded that the most effective parameter for Ra was the feed rate with 68.70% and the feed rate * cutting speed with 10.89%. It has been concluded that the least effective parameter is the depth of cut with 0.0006%. In the Anova analysis of the cutting force in Table 8, it was seen that the most effective parameters were the feed rate with 36.64% and the depth of cut with 18.65%, respectively. It has been concluded that the parameter that has the least effect on the cutting force is the depth of cut * feed rate with a ratio of 4.5%. The error occurred as a result of ANOVA analysis remained at 7.68% in surface roughness and 7.32% in cutting force. According to both surface roughness and shear force ANOVA results, it was determined that the most effective parameter was the feed rate.

V. CONCLUSION

In this study, surface milling process was performed on TOOLOX 44 material under dry cutting conditions, Taguchi method was used to determine the optimum machining conditions and ANOVA was used for statistical analysis results. The results of the study are listed below:

1. With the help of the graphs of the mean S / N ratios of the optimum levels of the control factors, the optimum parameter combinations for the surface roughness were found as the depth of cut 2nd level, the feed rate 1st level and the cutting speed 3rd level.
2. With the help of the graphs of the mean S / N ratios of the optimum levels of the control factors, the optimum parameter combinations for the cutting force were found as 1 for depth of cut, 1 for the feed rate and 4 for the cutting speed.
3. According to the results of the statistical analysis, it was concluded that the most effective factor (68.70%) on the surface roughness was the feed amount and the least effective (0.0006%) factor was the depth of cut.
4. According to the results of the statistical analysis, it was concluded that the most effective factor (36.64%) on the cutting force was the feed rate and the least effective (4.5%) factor was the depth of cut * feed rate.
5. It has been concluded that Taguchi method is an efficient method for the analysis and optimization of surface roughness and cutting force occurring in machining processes.

VI. ACKNOWLEDGEMENT

The data used in this study were taken from the Master's study titled " An Investigation into the Machinability of Hot Work Tool Steel (TOOLOX 44)" by Binali. Also, this work was supported by Scientific Research Projects Coordination Unit of Karabük University. Project Number: KBÜ-BAP-15/2-YL-043

REFERENCES

[1] T S Lee and Y J Lin, A 3D predictive cutting-force model for end milling of parts having sculptured surfaces, *Int J Adv Manuf Technol*, (16), pp 773-783, 2000

[2] E Şirin, Ş Şirin, Y Turgut ve İ Korkut, AISI D2 soğuk iş takım çeliğinin frezelenmesinde yüzey pürüzlülüğünün taguchi metodu ile optimizasyonu, *Düzce Üniversitesi Bilim ve Teknoloji Dergisi*, (3), pp 132-144, 2015

[3] J Z Zhang, J C Chen and E D Kirby, Surface roughness optimization in an end-milling operation using the Taguchi design method, *Journal of Materials Processing Technology*, 184(1-3), pp 233-239, 2007

[4] A Duran ve A Acır, HSS torna kalemindeki talaş açısının kesme kuvvetlerine etkisi, *Politeknik Dergisi*, 7(3), pp 211-215, 2004

[5] Ş Bayraktar ve Y Turgut, AL-5083 alaşımının frezelenmesinde kesme kuvveti, yüzey pürüzlülüğü ve çapak yüksekliğinin optimizasyonu, *7 th International Symposium on Machining*, İstanbul, 2016
[6] B İ Toprak, F M Çağlar, O Çolak, K Kıran ve M Bayhan, «Ti-6Al-4V süper alaşımının yüksek basınçlı soğutma kullanılarak frezelenmesinde yüzey pürüzlülüğünün Taguchi yöntemi ile optimizasyonu, *SDU International Technologic Science*, 4(2), pp 30-39, 2012
[7] S L Mike, C Joseoh and M Caleb, Surface roughness prediction technique for CNC end-milling, *Journal of Industry Technology*, 15(1), pp 1-6, 1998
[8] B Lahidi, Determining deflection for metal turning operation, *Journal Of Industrial Technology*, 13(2), pp 21-33, 1997
[9] G Meral, M Sarıkaya ve H Dilipak, «Delme işlemlerinde kesme parametrelerinin Taguchi yöntemiyle optimizasyonu, *Erciyes Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, 27(4), pp 332-338, 2011
[10] A Hasçalık and U Çaydaş, Optimization of turning parameters for surface roughness and tool life based on the Taguchi method, *The International Journal of Advanced Manufacturing Technology*, (38), pp 896-903, 2008
[11] T Ding, S Zhang, Y Wang and X Zhu, Empirical models and optimal cutting parameters for cutting forces and surface roughness in hard milling of AISI H13 steel, *The International Journal of Advanced Manufacturing Technology*, (51), pp 45-55, 2010
[12] B Gökçe ve S Taşgetiren, Kalite için deney tasarım, *Makine Teknolojileri Elektronik Dergisi*, 6(1), pp 71-83, 2009
[13] [Online] Available: https://www ssab com/products/brands/toolox/products/toolox-44?accordion=downloads
[14] S Neşeli ve E Türkeş, Tornalamada kesme parametrelerinin yüzey pürüzlülüğü üzerine olan etkilerinin taguchı metodu ile incelenmesi ve optimizasyonu, *16. Ulusal Makina Teorisi Sempozyumu*, Erzurum, 2013
[15] T Kıvak, Optimization of surface roughness and flank wear using the Taguchi method in milling of Hadfield steel with PVD and CVD coated inserts, *Measurement*, vol 50, pp 19-28, 2014
[16] K Palanikumar, Experimental investigation and optimisation in drilling of GFRP composites, *Measurement*, 44(10), pp 2138-2148, 2011
[17] N Masmiati andA A D Sarhan, Optimizing cutting parameters in inclined end milling for minimum surface residual stress–Taguchi approach, *Measurement*, vol 60, pp 267-275, 2015
[18] N Mandal, B Doloi, B Mondal and R Das, Optimization of flank wear using Zirconia Toughened Alumina (ZTA) cutting tool: Taguchi method and Regression analysis, *Measurement*, 44(10), pp 2149-2155, 2011
[19] S Neşeli, Tırlama titreşimleri üzerine süreç sönümleme etkisinin analitik olarak araştırılması ve tornalamada kararlı kesme derinliği ile süreç sönümleme değerlerine bağlı parametre optimizasyonu, Doktora Tezi, *Selçuk Üniversitesi Fen Bilimleri Enstitüsü*, 2013

Integration of exergy analysis and techno-economic optimization tool for the evaluation of hybrid renewable systems

MEHDI MEHRPOOYA[1], SEYED ALI MOUSAVI[1] and MOHAMMAD AMIN VAZIRI RAD [1]

[1] Department of renewable Energies and Environment, Faculty of New Sciences and Technologies, University of Tehran,Tehran, Iran, mehrpoya@ut.ac.ir

[1] Department of renewable Energies and Environment, Faculty of New Sciences and Technologies, University of Tehran,Tehran, Iran, s.a mousavi74@ut.ac.ir

[1] Department of renewable Energies and Environment, Faculty of New Sciences and Technologies, University of Tehran,Tehran, Iran, aminvaziririad@gmail.com

***Abstract* - The main goal of this project is to propose a powerful decision-making process to evaluate the hybrid renewable configurations in various applications. To this end, for the first time, the two methodologies of HOMER Pro software and exergy analysis are integrated. In this study, techno-economic analysis for the rural electrification purpose is investigated. Accordingly, several combined energy systems consist of Photovoltaic panels (PV), Wind Turbines (WT), Battery banks (Bat), and Diesel Generator (DG) are techno-economically investigated. Also, to improve technical evaluation of suggested hybrid renewable systems, the detailed exergy analysis is conducted for each component and the whole system. The results illustrated that the cost of energy and net present cost of best scenario are 0.17 $/kWh and 168378 $, respectively. The overall exergy efficiency and exergy destruction rate of optimum configuration (PV/WT/DG/Bat) are accounted for 25.96% and 303206.86 kWh/year, respectively. Also, it was found that 88.68% of the overall exergy wasting occurred in the PV modules. Correspondingly, to diminish the exergy wasting of the system, utilization of different common solar tracker is examined. To investigate the cost-effectiveness of each proposed tracking system, the Cost-Effective Index (CEI) is defined. The more the CEI, the more the cost-effectiveness of energy configuration. The vertical tracking system with a capacity of 20.1 kW PV has the greatest value of CEI (13.49), as a result, this system is the most cost-effective option. By employing this configuration, the overall exergy efficiency increases to 32.14%, and the overall irreversibility rate reduces to 224054.14 kWh/year. Finally, the parametric sensitivity analyses regarding the economic factors, exergy parameters, and renewable resources are conducted to generalize the outcomes of this project to other rural areas.**

Keywords* - *Rural Electrification; Hybrid Renewable System; Exergy Analysis; Economic Analysis; Solar Tracking;

I. INTRODUCTION

In recent years, most rural areas of Iran have been electrified, but there are still small villages in deprived cities that need the energy to develop and improve resident's quality of life [1]. In recent years, technico-economic optimization studies for such areas have been conducted using hybrid systems analysis tools. However, due to the reliance on software capabilities, the exergy aspect of hybrid systems has not been considered. Sinha and Chandel [2] compared different software tools and concluded that with highest usage rate in the field of hybrid renewable systems optimization, HOMER software is the most popular tool between researchers. They also stated that this software has suitable speed and accuracy compared to other techno-economic optimization methods. Li et al. [3] by using HOMER software designed a PV/WT/biogas/battery for power supply of a rural village in west China. Their results show that the proposed hybrid system with the installation capacities of 104 kW PV, 30 kW WT, 50 kW bio generator, 331 kWh battery, and 99 kW converter can economically make the village independent of the grid transmission lines by 0.201 $/kWh cost of energy. In another study, Ramesh and Saini [4] used of PV/WT/DG/Hydro turbine/battery to supply power in an Indian remote area. Their result showed that the proposed hybrid system can cover the electricity demand with about 0.162 $/kWh cost of energy. Also, they illustrated that for such stand-alone system using of lithium-ion batteries instead of lead-acid ones can reduce the system costs by about 30%. Johannsen et al. [5] investigated a PV-WT mini grid for rural electrification in Kenya. They showed that the combination of diesel generator with battery bank can reduce the cost of energy from 0.594 to 0.585 for a combined energy configuration with more than 99 renewable fractions. These results illustrate the importance of using a back-up energy source such as DG in hybrid renewable systems. H. Caliskan et al. [6] evaluated the exergetic performance of a combined renewable system based on power and hydrogen production. This hybrid system is composed of wind turbine, PV module, electrolyzer and PEMFC. It was found that the highest and least values of exergy efficiency belonged to the wind turbines (71.7%) and photovoltaic panels (9.67%), respectively. A detailed exergy evaluation and sensitivity analysis are conducted for a combined PV-wind system with hydrogen storage [7]. The gained findings illustrate that electrolyzer with 68.75% of exergetic efficiency has an acceptable performance. Also, the sensitivity of the exergy efficiency of the PV according to the variations of solar radiation is more than ambient temperature. Energy and exergy analyses are applied to a hybrid renewable system for hydrogen and power generation in island, Bozcaada [8]. The excess electricity

generated by renewable systems is used in electrolyzer for hydrogen production. It was denoted that about 35.48% of the input exergy (13848.18 kWh/day) is destroyed in PV panel. Also, electrolyzer unit has the greatest amounts of energy efficiency (59.68%) and exergy efficiency (60.26%). According to the presented literature reviews, the combination of exergy concept and HOMER Pro software could be a useful tool for the assessment of technical and economic aspects of hybrid renewable systems. In this investigation, for the first time, this novel approach is used to assess the hybrid renewable energy configurations for rural electrification. Based on the electricity demand and local renewable resources, several combined energy systems are simulated and proposed. Then, the optimum scenario is introduced. The exergy assessment is employed for each device and the entire hybrid system. Based on the obtained exergetic parameters, the outcomes of the exergy analysis indicated that for improving the system performance, the utilize of different tracking systems could be a suitable solution. Finally, to make the obtained outcomes generalized, the parametric analyses are conducted concerning some factors, e.g., the economic conditions, the exergetic indicators, the renewable component cost, and the local renewable resources potential.

II. METHODOLOGY

Ting (25°16′24″N & 61°06′47″E) is a remote village located in the Sistan & Baluchestan province. This rural area has no access to the grid. Accordingly, the utilization of the combined renewable system can be considered as one of the effective solutions for electricity supply. Figure 1 demonstrates the monthly average solar radiation and clearness index for Ting village. This rural area with average annual radiation of 5.53 kWh/m^2/day has appropriate solar energy potential. The largest amount of radiation is about 7.12 kWh/m^2/day in June and the lowest radiation is about 3.49 kWh/m^2/day in December. Figure 2 indicates the monthly average wind velocity and ambient temperature for Ting village. The highest and least wind speed are associated with June (6.76 m/s) and November (5.09 m/s), respectively. The yearly mean wind velocity is 6.05 m/s. The maximum and minimum values of ambient temperature occurred in the July (31.88 °C) and January (10.42 °C), respectively. The mean ambient temperature of this area is 22.23 °C. So, there is no undesirable temperature effect for the PV modules.

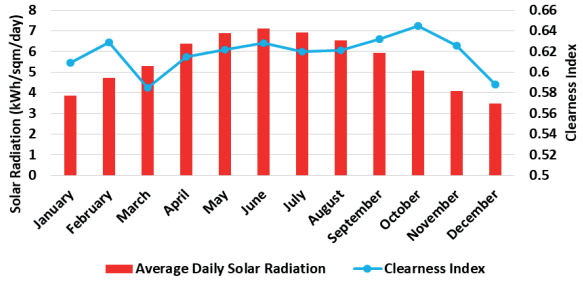


Figure 1. Average solar radiation and clearness index for the selected area.

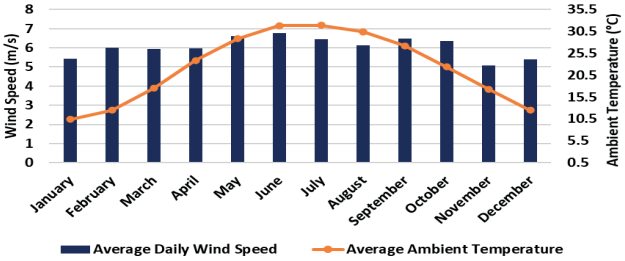


Figure 2. Average wind speed and ambient temperature for the selected area.

The main financial factors provided by this software are introduced below:

The Net Present Cost is considered as the present value of all the costs that it incurs over the lifespan of the project, minus the salvage cost at the end of the project lifespan. This economic indicator can be expressed as follows:

$$NPC = \frac{C_{ann.tot}}{CRF(i,n)} \quad (1)$$

In which, $C_{ann.tot}$ ($/year) implies to the overall annualized cost. CRF is the Capital Recovery Factor which could be computed as follows:

$$CRF(i,n) = \frac{i(1+i)^n}{(1+i)^n - 1} \quad (2)$$

In the above definition, n and i (%) are related to the project lifespan and annual real interest rate, respectively.

The annual real interest rate depends on inflation rate f (%) and nominal interest rate i' (%) which could be formulated as follows:

$$i = \frac{i' - f}{1 + f} \quad (3)$$

In this research, the lifespan of the project is considered to be 20 years, as well as, the inflation rate and the discount rate are 15% and 18%, respectively [9].

The average cost per useful power produced by an energy configuration is defined as the Cost of Energy (COE). This index is utilized for evaluating the cost-effectiveness of the energy configuration and can be determined as below:

$$COE = \frac{C_{ann.tot}}{E_{served}} \quad (4)$$

In which E_{served} (kWh/year) represents the annual energy generated.

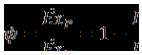
The technical and economic specifications of the components are available in Table 1.

Table 1. The properties of the components utilized in simulations [10].

Generator model	Generic Small Genset (size-your-own)
Diesel fuel price ($/L)	0.1
Minimum load ratio (%)	25
Generator lifespan (h)	15000
Capital cost ($/kW)	450
Replacement cost ($/kW)	425
O&M cost ($/h)	0.02
Turbine model	AWS HC 1.5kW
Power coefficient (C_P)	0.45
Turbine lifespan (year)	20
Cut-in speed (m/s)	2.2
Cut-out speed (m/s)	12

Capital cost ($/unit)	3000
Replacement cost ($/unit)	2700
	30
PV model	Sharp ND-250QCS
	15.3
	Flat plate
Temperature coefficient (%/ C)	-0.485
Panel lifespan (year)	20
NOCT (C)	47.5
Capital cost ($/kW)	1300
Replacement cost ($/kW)	1300
	1% Capital

The main target of the exergy-based concept is to identify weak points of the thermodynamic systems where the devices do not effectively consume energy and waste it. The second stage after specifying the components with the highest values of irreversibility is finding the solutions for reducing the exergy wasting of the system. By considering control volume around each component and the entire system, the exergetic outputs, such as exergy efficiency, irreversibility rate, and sustainability index are computed. For each unit, the exergy efficiency and destroyed exergy can be formulated as below:



(5)

$$\dot{E}x_D = \dot{E}x_F - \dot{E}x_P - \dot{E}x_L \quad (6)$$

The total exergy efficiency and exergy destruction rate for this combined energy system can be formulated as follows:

$$\psi_{Overall} = \frac{P_{WT}+P_{PV}+P_{DG}}{\dot{E}x_{ch,fuel}+\dot{E}x_{Sun}+(\dot{E}x_{K,in}-\dot{E}x_{K,out})} \quad (7)$$

$$\dot{E}x_{D,Overall} = \dot{E}x_{D,WT} + \dot{E}x_{D,PV} + \dot{E}x_{D,DG} + \dot{E}x_{D,Conv} \quad (8)$$

With the application of each type of tracker, the overall cost of the system increases. So, it must be determined which tracking method increases the exergy efficiency more than the price increase. This techno-economic indicator, which is the result of the integration of exergy analysis and HOMER Pro software, could be formulated as follows:

$$CEI = \frac{\%\Delta\psi_{Overall}}{\%\Delta COE} \quad (9)$$

It can be noticed that CEI is defined as the ratio of the percentage difference of exergy efficiency to the percentage difference of COE.

III. RESULTS AND DISCUSSIONS

The hybrid system comprised of 26.4 kW PV, 8 units WT, 5 kW DG, and 91 units Bat is the optimum option. The COE and NPC of this scenario are 0.17 $/kWh and 168378 $, respectively. Due to the application of the DG as the backup unit, the Renewable Fraction (RF) is not 100%, and the amount of 18831 kg/year CO_2 is emitted. Now, the technical performance of this hybrid system must be evaluated by the exergy method. Table 8 represents the results of the exergy analysis. The exergetic calculations illustrated that the total exergy efficiency and exergy destruction rate of this configuration are 25.96% and 303206.86 kWh/year, respectively. So, the main origin of exergy wasting must be determined. Table 2 illustrates the exergetic parameters for the suggested hybrid systems.

Table 2. The exergetic parameters for suggested energy scenarios.

Scenario	*Overall Exergy Efficiency (%)*	*Annual Exergy Destruction (kWh/year)*
	25.96	303206.86
Vertiacl (C)	27.93	286837.83
Horizantal (C)	27.03	295401.59
2 Axes (C)	29.12	278939.20
Vertiacl (O)	32.14	224054.14
Horizantal (O)	28.64	250500.43
2 Axes (O)	34.38	197499.02

Figure 3 exhibits the pie chart for exergy destruction rate of the optimum system. It can be seen that the highest share in the total irreversibility rate is associated with PV modules (88.68%), followed by the wind turbines (7.83%). So, the photovoltaic panels have the most outstanding role in the exergy wasting of the system. By improving the operating conditions of PV modules, the performance of the overall system is considerably improved. In this investigation, to improve the exergetic parameters (technical aspect) of the hybrid energy system, the use of three types of tracking systems is recommended.

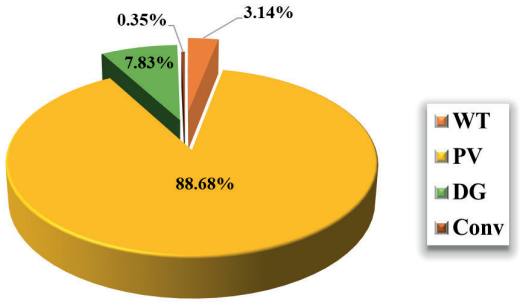


Figure 3. The contribution of components to the overall exergy destruction for the optimum scenario.

Figure 4 states the influence of tracking systems on PV production and the exergy destruction rate. At a constant PV capacity (26.4 kW), with the utilize of horizontal, vertical, and dual axes trackers, the PV electricity production increases, by 7.79%, 21.37%, and 32.18%, respectively. Also, the highest reduction in exergy wasting is related to the two axes tracker (247599.48 kWh/year). Amongst the all suggested modes, the lowest value of the irreversibility rate belonged to the dual axes tracker with a PV capacity of 17.8 kW. It can be deduced that dual axes system is the most efficient option, followed by the vertical tracker. Figure 5 exhibits the relationship between COE and the exergy efficiency of different tracking modes. It can be shown that at the constant capacity of PV modules (26.4 kW), the dual axes tracking system has the greatest exergy efficiency (29.12%), but by using this tracking system, the COE increases by 12.35% compared to the optimum scenario. On the other hand, when vertical tracking is utilized, with an increase of 5.88% of COE, the exergetic efficiency increases to 27.93%. So, the cost-effectiveness of suggested tracking modes must be evaluated.

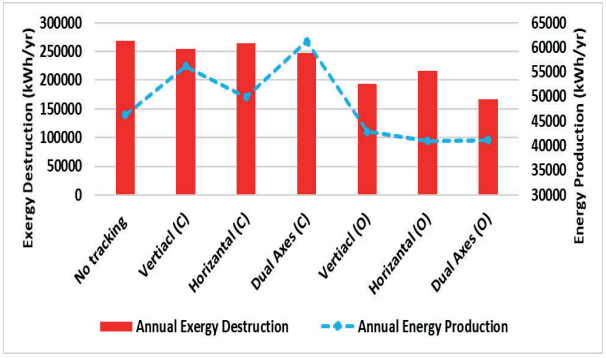


Figure 4. The mounts of PV energy production and exergy destruction in different tracking methods.

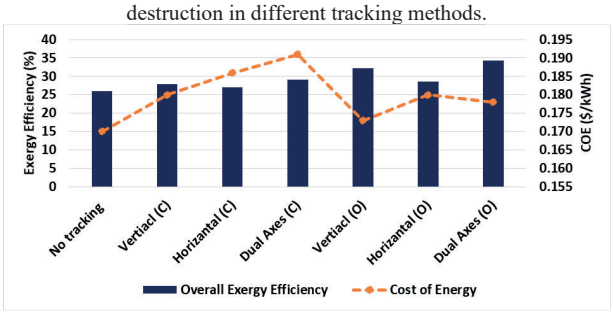
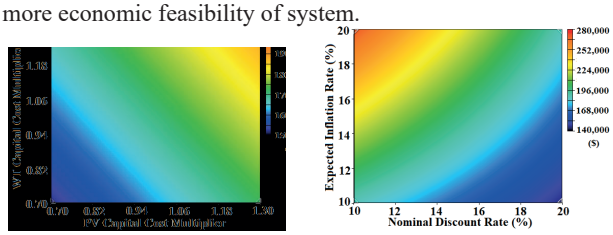


Figure 5. The exergy efficiency and COE of the different tracking methods.

Figure 6 (a) demonstrates the impact of alternation in the PV and WT capital costs on the overall NPC. It can be shown that with a 30% increase in capital costs of PV and WT, the overall NPC increases to 190283 $. At the same time, a 30% decrease in both costs causes the overall NPC reduces to 152286 $. Accordingly, the capital cost of the renewable components can be considered a crucial factor for investors. Figure 6 (b) represents the sensitivity heat map regarding the impact of nominal discount rate and expected inflation rate on overall NPC. Actually, at a certain amount of the nominal discount rate, by increasing the inflation rate, the NPC increases which consequently will decrease the cost-effectiveness of the use of vertical tracking system. In fact, lower inflation rate leads to more economic feasibility of system.



a) Wind turbine capital cost Vs. PV capital cost b) Inflation rate Vs. nominal discount rate

1

Figure 6. NPC sensitivity analysis for the optimum scenario in vertical tracking mode.

Figure 7 represents the magnitudes of CEI for different tracking systems. As the value of CEI increases, the cost-effectiveness of the hybrid system increases. It can be shown that the highest value of CEI is associated with the vertical tracking mode with a capacity of 20.1 kW PV (13.49). Consequently, this configuration is the most cost-efficient alterative option for the optimal scenario. Although the greatest exergetic efficiency belonged to the Dual Axes (O), but this configuration is the second cost-effective option (CEI=6.89). Also, the amounts of CEI for the suggested horizontal tracking systems are low, which indicates these scenarios do not significantly increase exergy efficiency relative to the cost of the overall system. Based on the figure 20, the horizontal tracking system is not cost-effective neither as constant solar capacity nor as optimizer mode.

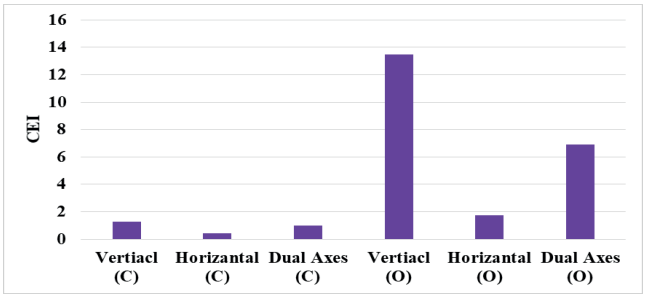


Figure 7. The values of CEI for different solar tracking modes.

IV. CONCLUSIONS

This investigation attempts to determine a cost-effective hybrid energy system for a remote village. For implementing the comprehensive techno-economic evaluation for the proposed energy configurations, the exergy analysis is linked to the HOMER Pro simulator. The most noticeable findings of this project can be categorized as follows:

- In order to provide the electricity demand of Ting village (176 kWh/day), the hybrid energy system comprised of 26.4 kW PV, 8 units of WT, 5 kW DG, and 91 units of Bat is chosen as the optimal scenario. The amounts of NPC and COE for this scenario are calculated to be 168378 $ and 0.173 $/kWh, respectively.
- Based on the performance analysis, this combined system has an exergetic efficiency of 25.96% and an irreversibility rate of 303206.86 kWh/year. Also, the largest contribution to the overall exergy wasting occurred in the PV modules (88.68%), followed by wind turbines (7.83%).
- Technically, the dual – axis tracking system composed of 17.8 kW PV, 10 units of WT, 5 kW DG, and 85 units of Bat is determined as the most efficient configuration. The overall exergy efficiency and irreversibility rate of this hybrid configuration are computed to be 34.38% and 197499.02 kWh/year, respectively.
- The highest value of CEI (13.49) is related to the vertical tracking system with a capacity of 20.1 kW PV. Therefore, this configuration is the most cost-effective option compared to the no tracking system. With the application of the vertical tracker, the exergy destruction of PV reduces by 27.77%, and the mean sustainability index increases by 2.65%. The least values of CEI belonged to the horizontal tracking systems. Actually, the horizontal trackers showed the lowest economic efficiency.

References

1. Ghasemi, A., et al., *Techno-economic analysis of stand-alone hybrid photovoltaic–diesel–battery systems for rural electrification in eastern part of Iran—A step toward sustainable rural development.* Renewable and Sustainable Energy Reviews, 2013. **28**: p. 456-462.
2. Sinha, S. and S. Chandel, *Review of software tools for hybrid renewable energy systems.* Renewable and Sustainable Energy Reviews, 2014. **32**: p. 192-205.
3. Li, J., P. Liu, and Z. Li, *Optimal design and techno-economic analysis of a solar-wind-biomass off-grid hybrid power system for remote rural electrification: A case study of west China.* Energy, 2020. **208**: p. 118387.
4. Ramesh, M. and R.P. Saini, *Dispatch strategies based performance analysis of a hybrid renewable energy system for a remote rural area in India.* Journal of Cleaner Production, 2020: p. 120697.
5. Johannsen, R.M., P.A. Østergaard, and R. Hanlin, *Hybrid photovoltaic and wind mini-grids in Kenya: Techno-economic assessment and barriers to diffusion.* Energy for Sustainable Development, 2020. **54**: p. 111-126.
6. Caliskan, H., I. Dincer, and A. Hepbasli, *Energy, exergy and sustainability analyses of hybrid renewable energy based hydrogen and electricity production and storage systems: modeling and case study.* Applied Thermal Engineering, 2013. **61**(2): p. 784-798.
7. Calderón, M., et al., *Evaluation of a hybrid photovoltaic-wind system with hydrogen storage performance using exergy analysis.* International journal of hydrogen energy, 2011. **36**(10): p. 5751-5762.
8. Kalinci, Y., I. Dincer, and A. Hepbasli, *Energy and exergy analyses of a hybrid hydrogen energy system: A case study for Bozcaada.* international journal of hydrogen energy, 2017. **42**(4): p. 2492-2503.
9. Mousavi, S.A., et al., *Decision-making between renewable energy configurations and grid extension to simultaneously supply electrical power and fresh water in remote villages for five different climate zones.* Journal of Cleaner Production, 2020: p. 123617.
10. Razmjoo, A., et al., *A Technical analysis investigating energy sustainability utilizing reliable renewable energy sources to reduce CO2 emissions in a high potential area.* Renewable Energy, 2020.

Simulation and Optimization of an Automotive PEM Fuel Cell System by using EES Software and Genetic Algorithm

Fathollah Pourfayaz[*,1], Arman Ekrataleshian[1], Mehdi Mehrpooya[1]
[1] Department of Renewable Energies and Environment, Faculty of New Sciences and Technologies, University of Tehran, Tehran/Iran, pourfayaz@ut.ac.ir (Corresponding Author)
[1] Department of Renewable Energies and Environment, Faculty of New Sciences and Technologies, University of Tehran, Tehran/Iran, ekrataleshian.ar@ ut.ac.ir
[1] Department of Renewable Energies and Environment, Faculty of New Sciences and Technologies, University of Tehran, Tehran/Iran, Tehran/Iran, mehrpoya@ut.ac.ir

Abstract **- Fuel cells are electrochemical energy conversion devices that offer a high-efficiency and low-emission alternative to internal combustion engines (ICE) in-vehicle applications. Among the types of fuel cells, the polymer electrolyte membrane (PEM) fuel cells are particularly suitable for use in passenger vehicles, such as cars and buses due to their fast startup time, low operating temperatures, and favorable power-to-weight ratio. In this study, a PEM fuel cell has been modeled by using Engineering Equation Solver (EES) software and then has been optimized using the Genetic algorithm-based optimization method. In order to maximize the energy efficiency, the influences of the electrode area (cross-sectional) of PEM, the thickness of PEM, the cell width, the number of cells in series per stack, the total current of the stack, and operating temperature as independent variables on the energy efficiency as the objective function have been investigated. The results showed the energy efficiency increases by increasing the electrode area (cross-sectional) of PEM and the number of cells in series per stack and decreases by increasing the thickness of PEM, cell width, and the total current of the stack. Moreover, at optimal values of independent variables, the energy efficiency can be reached about 54%.**

Keywords **- automotive, PEM fuel cells, Genetic Algorithm, Energy efficiency**

I. Introduction

FUEL cells are kinds of electrochemical energy conversion devices, which convert fuel's chemical energy to power. These devices are more efficient than internal combustion engines (ICEs) in converting fuel's energy to power. Furthermore, the emissions of them are fewer than ICEs. Therefore, they have a high potential to be a good replacement for ICEs in vehicles [1].

In a fuel cell, fuel (e.g., hydrogen gas) and the oxidant (e.g., oxygen gas from the air) are used to generate electricity, while heat and water are typical by-products of the fuel cell operation. A fuel cell typically works on the following principle: as the hydrogen gas flows into the fuel cell on the anode side, a platinum catalyst facilitates oxidation of the hydrogen gas which produces protons (hydrogen ions) and electrons. The hydrogen ions diffuse through a membrane (the center of the fuel cell which separates [2, 3].

Today, several in-plant aircraft, automobiles, spacecraft and similar vehicles use efficient fuel cells. Between different types of fuel cells, the polymer electrolyte membrane fuel cells (PEMFC) have shown special characteristics for employing in transportation, portable, and stationary applications. This system has low emission, high efficiency, low (60-90℃) operating temperature, and has no harmful by-products [4, 5].

According to what is mentioned, since polymer electrolyte membrane fuel cells efficiently can produce high power density, therefore this technology is potentially attractive for employing in automotive applications. Unlike residential and stationary fuel cells, automotive PEMFCs undergo the entire slew of aggressive variable loads and environmental conditions that are typically experienced by conventional ICEs. Most of the automotive companies started their research and development (R&D) on PEMFC between 1998 and 2010.

The performance, durability, and cost of the PEMFC system and stacks for automotive are approaching levels such that major automotive manufacturers should acquire some specific targets for commercialization. In order to achieve commercialization, the targets of performance, durability, and cost of the fuel cell components (catalyst, membrane, diffusion media, bipolar plates, etc.,) and the system (thermal, fuel, air, and water management, and balance of plant) have to be met simultaneously. Therefore the performance of the system is one of the important subjects that researchers should focus on that [6].

As mentioned earlier, one of the advantages of PEMFCs is their low operating temperature. This parameter usually is controlled by an external cooling source. Since the performance of the cell is related and sensitive to its temperature, many studies are dedicated to surveying this parameter. Ferng et al. [7] presented an analytical and experimental study of a PEMFC to analyze the effect of operating pressure and temperature on the performance of the cell. The results demonstrate the increment in cell performance at elevated temperatures. The predicted performance curves are in agreement with the experimental data. In another study, Santarelli et al. [8] performed an experimental analysis to investigate the influence of some operating variables on the performance of a single

PEMFC. It is concluded that an increase in the reactant saturation temperature improves the performance of the cell, especially in low and medium loads.

To deliver the desired power by PEMFCs are combined in parallel or series or a combination of both ways to create a stack. With an increase in cell surface area, each cell will produce more current. The generated current affects the energy efficiency of the fuel cell. Another effective parameter on the performance of the system is the thickness of membranes that is investigated in the literature.

Therefore for obtaining a high-performance PEMFC stack the aforementioned parameters should be investigated. In this study polymer electrolyte membrane fuel cell is simulated in EES software. After that, by employing Genetic algorithm optimization, the optimum values of design parameters for maximizing the energy efficiency are obtained.

II METHODOLOGY

A. Assumption

The following assumptions can be made about the water-splitting reaction: (a) Hydrogen and oxygen are ideal gases, (b) water is an incompressible fluid, and (c) the gas and liquid phases are separate. Based on these assumptions the change in enthalpy of the water-splitting reaction can be calculated regarding pure hydrogen (H_2), oxygen (O_2), and water (H_2O) at a standard temperature and pressure (25°C and 1 atm). The total change in enthalpy for splitting water is the enthalpy difference between the products and the reactants.

B. Thermodynamic modeling

PEMFCs consists of three main parts: Anode, cathode, and an electrolyte that is a solid, thin polymeric sheet. The various reactions for a PEMFC fed with a hydrogen-containing anode gas and an oxygen-containing cathode gas, are:

$$Anode: H_2 \Leftrightarrow 2H^+ + 2e^- \quad (1)$$

$$Cathode: 2H^+ + 2e^- + 0.5O_2 \Leftrightarrow H_2O \quad (2)$$

$$Overall: H_2 + 0.5O_2 \Leftrightarrow H_2O \quad (3)$$

The above equations generate water, heat, and electricity. The following equation expresses the voltage of cell [9]:

$$V_{cell} = E_{Nerst} + \eta_{act} + \eta_{ohmic}$$

Where E_{Nerst}, η_{act} and η_{ohmic} are thermodynamic potential, the activation overvoltage, and the ohmic overvoltage. The three overvoltage terms are all negative and show reductions from E_{Nerst} to give the useful cell voltage.

The thermodynamic potential, the cathode activation overvoltage, the anode activation overvoltage, and the ohmic overvoltage can be obtained as below [10]:

$$E_{Nerst} = 1.23 - 8.5 * (T_{stack} - 298) + 4.31e^{-5} * T_{stack} * \ln(p_{H_2} + 0.5 p_{O_2}) \quad (4)$$

$$\eta_{act} = -0.95 + 2.43e^{-3} * T_{stack} + 1.92e^{-4} * T_{stack} * \ln(A_{PEM}) - 1.92e^{-4} * T_{stack} * \ln(I_{FC}) + 7.6e^{-4} * T_{stack} * \ln(c_{O_2}) \quad (5)$$

$$\eta_{ohmic} = ((-I_{FC} * t_{PEM}) / A_{PEM}) * (8 / \exp(3.6 (T_{stack} - 353) / T_{stack})) * (1 + 1.64 * I_{FC} / A_{PEM} + \gamma * (I_{FC} / A_{PEM})^3 \quad (6)$$

The total amount of energy released in the fuel cell reaction is equivalent to the change in enthalpy. The standard enthalpy for splitting water is 286 kJ mol^{-1}. The total energy demand is related to the thermoneutral cell voltage by the expression:

$$V_{tn} = \Delta h / n * F$$

Where n and F are electronic charge number and faraday constant. Δh is the variation in enthalpy of the overall equation.

Finally, the energy efficiency of the system can be shown by the following equation:

$$\eta_{energy} = V_{cell} / V_{tn}$$

III. RESULTS AND DISCUSSION

A. Sensitivity Analysis

In this section, by employing some design variables, the variation of energy efficiency is investigated. In Figure 1 the variation of energy efficiency versus axis electrode area of PEM is shown. As can be seen, by increment of the electrode area of the fuel cell, the energy efficiency faces increment.

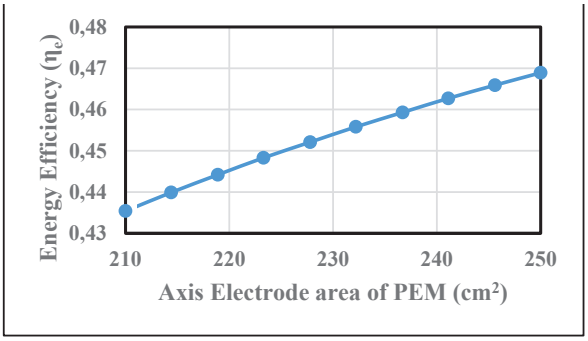


Figure 1. Variation of energy efficiency versus electrode area

Figure 2 depicts the changes in energy efficiency by an increase in the thickness of PEM. The results revealed that the increment of this parameter leads to a decrease in energy efficiency.

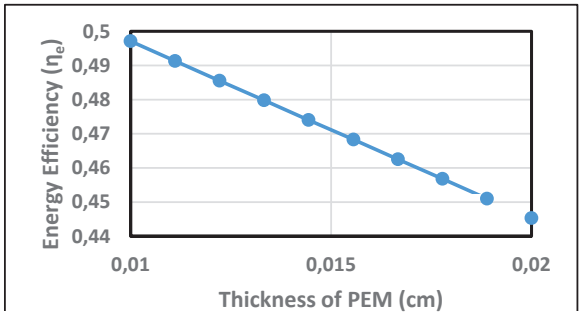


Figure 2. Variation of energy efficiency versus thickness of PEM

The variation of energy efficiency versus the cell width is plotted in Figure 3. As can be seen, cell width has a minor effect on energy efficiency and by the increment of this parameter, energy efficiency faces a slight decrement.

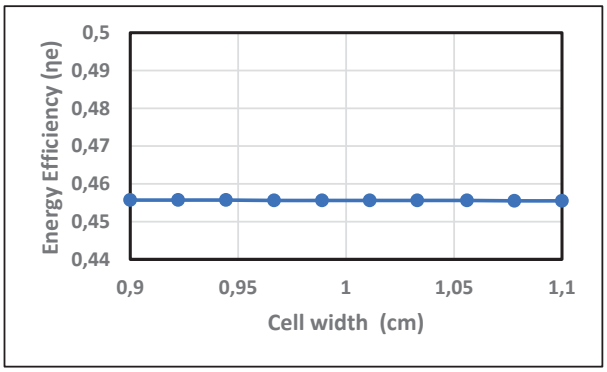


Figure 3. Variation of energy efficiency versus cell width

Figure 4 demonstrates the effect of variation of the number of cells in series per stack on energy efficiency. It is concluded that an increase in this parameter has a slightly positive effect on energy efficiency. By increment of the number of cells in series per stack, energy efficiency faces a little bit increment.

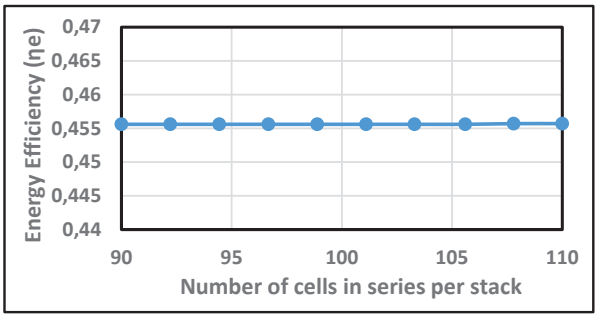


Figure 4. Variation of energy efficiency versus the number of cells in series per stack

Figure 5 shows the change in energy efficiency by variation of the operating temperature. The results reveal that by an increase in operating temperature from 333 K to 355 K, the energy efficiency has an ascending trend and increases.

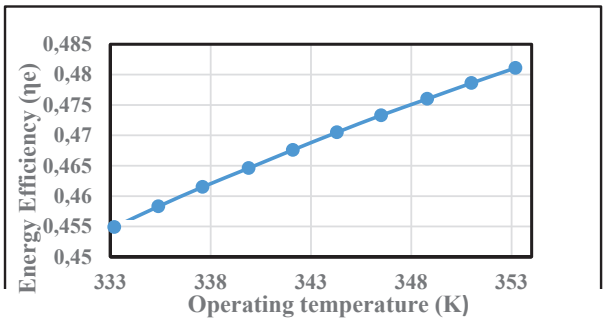


Figure 5. Variation of energy efficiency versus the operating temperature

In Figure 6 the effect of the total current of the stack on energy efficiency is surveyed. The results depict that by increasing the total current of the stack, the energy efficiency decreases.

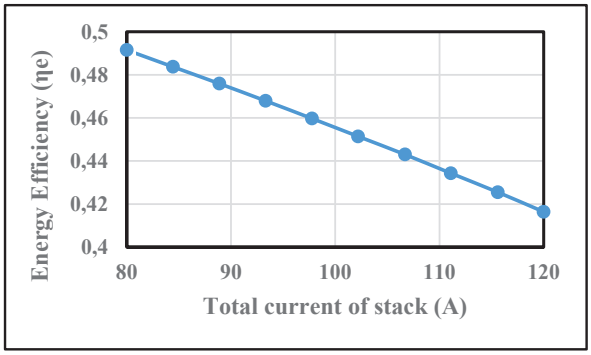


Figure 6. Variation of energy efficiency versus the total current of the stack

B. Optimization results

In this section by an objective function, which is energy efficiency is optimized. For this purpose, 6 independent variables including the number of cells in series per stack, electrode area (cross-sectional) of PEM, the thickness of PEM, cell width, operating temperature, and total current of the stack are considered. The variation range of these parameters is presented in Table 1.

Table 1. Range of independent variables

Independent Variable	Range
number of cells in series per stack	90 – 110
electrode area of PEM(cm2)	210-250
thickness of PEM(cm)	0.01-0.02
cell width(cm)	0.9-1.1
operating temperature(K)	333.2-353.2
the total current of the stack (A)	80-120

After optimizing the maximum value of energy efficiency is reached to 54.24 %. The optimum values of independent variables are presented in Table 2.

Table 2. Optimization results

Independent Variable	Range
number of cells in series per stack	100.5 ≈ 100
electrode area of PEM(cm2)	48.6
thickness of PEM(cm)	0.01016
cell width(cm)	1.064
operating temperature(K)	353.2
the total current of the stack (A)	80.18
Optimization Results	
Max. Value Energy Efficiency = 54.24 %	

IV. Conclusion

Simulation and optimization of an automotive PEM Fuel Cell System by using the EES Software and Genetic Algorithm are surveyed in this study. A sensitivity analysis is applied to investigate the effect of some parameters such as axis electrode area, the thickness of PEM, cell width, number of cells in series per stack, operating temperature, and total current of the stack on the energy efficiency. It is shown that increment of the axis electrode area, operating temperature, and the number of cells in series per stack lead to an increase in energy efficiency. On the other hand, by increasing the total current of the stack, thickness of PEM, and cell width, the energy efficiency decreases. Moreover, the result of optimization by the Genetic Algorithm is revealed that at the optimal point of independent variables, the energy efficiency can be increased to 54.24 %.

References

1 Larminie, J , A Dicks, and M S McDonald, *Fuel cell systems explained* Vol 2 2003: J Wiley Chichester, UK

2 Grujicic, M and K M Chittajallu, *Design and optimization of polymer electrolyte membrane (PEM) fuel cells.* Applied Surface Science, 2004 **227**(1): p 56-72

3 Chakraborty, U K , T E Abbott, and S K Das, *PEM fuel cell modeling using differential evolution.* Energy, 2012 **40**(1): p 387-399

4 Wu, B , M Matian, and G J Offer, *Hydrogen PEMFC system for automotive applications.* International Journal of Low-Carbon Technologies, 2012 **7**(1): p 28-37

5 Sharaf, O and M Orhan, *An overview of fuel cell technology: Fundamentals and applications.* Renewable and Sustainable Energy Reviews, 2014 **32**: p 810-853

6 Kocha, S S , *Polymer electrolyte membrane (PEM) fuel cells, automotive applications*, in *Fuel Cells* 2013, Springer p 473-518

7 Ferng, Y M , et al , *Analytical and experimental investigations of a proton exchange membrane fuel cell.* International Journal of Hydrogen Energy, 2004 **29**(4): p 381-391

8 Santarelli, M G and M F Torchio, *Experimental analysis of the effects of the operating variables on the performance of a single PEMFC.* Energy Conversion and Management, 2007 **48**(1): p 40-51

9 Ulleberg Ø (2001) Evaluation of IFE's 100 W PEM Fuel Cell Stack Performance Internal report, I f E T

10 Amphlett, J C , et al , *A model predicting transient responses of proton exchange membrane fuel cells.* Journal of Power sources, 1996 **61**(1-2): p 183-188

A Finite Element Based Methodology for Frictional Rolling Contact Problems

E.E. ÜSTÜN[1], K.B. YILMAZ[2], E. ACAR[1], and M.A. GÜLER[3]
[1] TOBB University of Economics and Technology, Ankara/Turkey, eustun@etu.edu.tr
[1] TOBB University of Economics and Technology, Ankara/Turkey, acar@etu.edu.tr
[2]Hacettepe University, Ankara/Turkey, korhanb.yilmaz@hacettepe.edu.tr
[3]American University of the Middle East, College of Engineering and Technology, Kuwait, mehmet.guler@aum.edu.kw

Abstract - **In this study, we develop a finite element (FE) based methodology to solve the frictional rolling contact problem between a rolling cylinder and a half-space. The developed methodology can be applied to solve the frictional rolling contact problems for cylindrical roller bearing problem. A three-dimensional (3D) explicit FE model is developed to simulate the rolling contact problem, and the contact stresses and contact widths are determined. It is assumed the half-plane is linearly elastic and static/kinematic coefficients of friction are constant throughout the analysis. The FE study focuses on determination of the contact stresses as a function of the value of external load and coefficient of friction. The results regarding the contact stresses, the in-plane component of the surface stresses and stick/slip distributions are presented. Various stick/slip distributions depending on different tractions applied to the rolling cylinder are determined though FE analyses. The FEA results are validated through analytical formulations. Finally, the roller-raceway interaction in the cylindrical roller bearing problems are used to demonstrate the developed methodology.**

Keywords - **cylindrical bearing, finite element method, normal contact solution, rolling contact, tangential contact solution**

I. INTRODUCTION

Rolling elements carry a significant role in many engineering systems; consequently, the discipline of rolling contact mechanics is involved in various fields. Road and railway dynamics, mechanical joints, mechanical systems with rolling elements are just few of the areas. Ball or roller bearings are mechanical components for which the rolling contact problem plays a major role. For these components, experimental determination of the contact stresses and patches is not possible in many circumstances, particularly in dynamic conditions. Hence, the examination methods are generally established on analytical or computer-based approaches.

Conventionally, the rolling contact problem is divided into two parts: normal and tangential problems. While the normal solution is related to the normal pressure and the contact patch; the tangential problem is associated with friction and its effects. The surface shear stress distribution depends on the traction force and coefficient of friction. The distribution of surface shear stress is expressed with the slip and stick areas in the contact patch. Slip is the relative velocity between the two material particles that are in contact, and its size in the contact patch changes depending on the value of traction.

Hertz solved the normal problem of contact between quadratic surfaces in 1882 [1]. For 2D frictional rolling contact problem, Carter developed his model and handled the surface shear stress in the contact patch [2]. Vermeulen and Johnson asserts a model for 3D frictional rolling contact by assuming an elliptical contact area [3]. Then, Kalker solved the 3D frictional rolling contact problem for arbitrary creepages in both lateral and longitudinal directions using several numerical methods [4]. Alinia and Guler put forward a 2D analytical solution for the problem, which can be applied to 3D problems with its adaptivity for both plane stress and plane strain geometries [5]. In this study, Guler et al. (2012) analytical formulation is used as a basis for validation of the Finite Element (FE) results. As in aforementioned solution methods, this formulation is also based on linear-elasticity, statics and half-space assumption.

In the FE modeling part, studies focusing on the tangential solution is still rare, although some static approaches [6-7] or 2D explicit solutions exist [8]. Zhao and Li propose an implicit-explicit sequential solution method for the 3D transient problem in wheel-rail interactions [9-11].

Differently from the method of Zhao and Li [9-11], this study proposes a one-step FE methodology for the 3D frictional rolling contact problems. Contact stresses, surface shear stresses, slip/stick distributions according to the traction coefficients are determined by the FE study. The findings are validated by the analytical formulation of Alina and Guler [5].

II. THE FINITE ELEMENT MODEL

A schematic diagram of the frictional rolling problem of our interest is shown in Fig. 1. The roller has a cylindrical contact surface, while the half-space has a flat one without curvature. Both contact surfaces are deformable. The Augmented Lagrange contact formulation is used in the FE analyzes, since the Augmented Lagrange formulation properly detects the discontinuities in the contact zone and has the ability to stabilize the stiffness matrix in each iteration throughout the solution process [12]. The element types selected in accordance with rolling contact modeling are 8-node CONTA174 and TARGE170. TARGE170 is used for roller's contact surface and CONTA174 for half-space's contact surface. These element types are frequently used in the contact of two deformable elastic surfaces that support 3-D contact in ANSYS/LS-DYNA

[12]. Throughout the FE model, SOLID185 type 8-node and 3 degree of freedom elements are preferred.

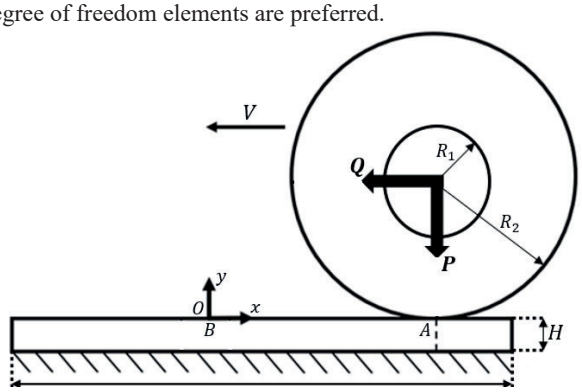


Figure 1: Schematic diagram of the 3D transient FE model.

The material behavior is defined as homogenous, isotropic, and linear elastic for all elements in the FE model. Structural steel is used for both roller and half-space. The material parameters are given in Table 1.

Table 1: Material parameters.

Young's modulus, E	210 GPa
Poisson's ratio, ν	0.3
Density, ρ	7850 kg/m^3

Table 2 lists the geometric parameters of the roller and half-space. Roller has an outer radius of 30 mm and inner radius of 10 mm, with a thickness of 1 mm. Half-space has the same thickness value with the roller, a height of 3 mm and a width of 55 mm.

Table 2: Geometric parameters (mm).

Inner radius of roller, R_1	10
Outer radius of roller, R_2	30
Height of half-space, H	3
Width of half-space, W	55
Thickness, t	1

The half-space is divided into two bodies to provide computational efficiency in the solution. In the contact surface of the upper half-space, a dense and uniform mesh is employed. The element edge size in the surface is 0.05 mm. Roller part is also divided into bodies for both the ease of boundary condition application and meshing. One quarter of the roller is finely meshed, as seen in Figure 2 and 3. The element edge size in contact surface of the roller is 0.1 mm. Total number of nodes are 147,813 and total number of elements are 114,254. Each element in the system is hexagonal and the system is uniformly meshed.

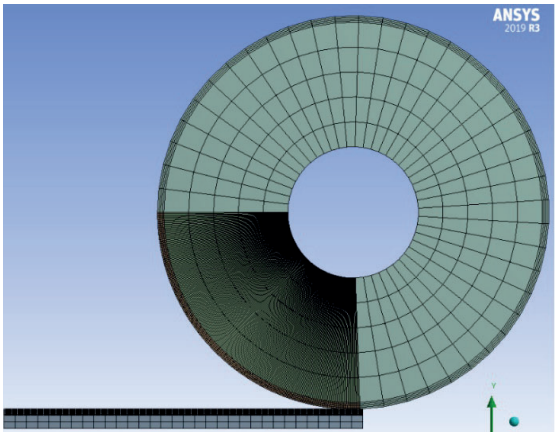


Figure 2: FE mesh of the model

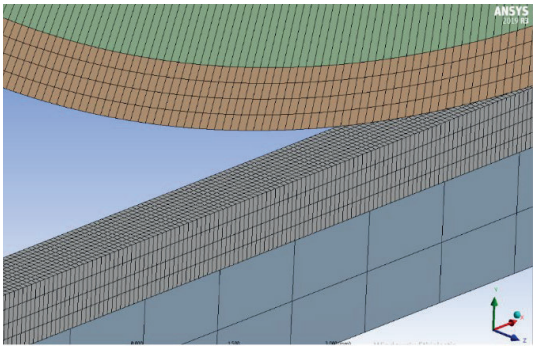


Figure 3: Zoomed-in view to contact surfaces

The bottom face of the half-space is fixed. The remote displacement condition is assigned to the inner and outer cylindrical surfaces of the gray region of the roller, which is shown in Figure 4. This condition allows the roller to move in x, y directions and rotate about the z-axis. Other degrees of freedom of the roller are constrained. Since the remote displacement condition in ANSYS brings a rigid behavior where it is defined, it is applied to the non-contact cylindrical surfaces. The outer cylindrical surfaces of the green region have the deformable behavior, the deformable-to-deformable contact hereby is provided.

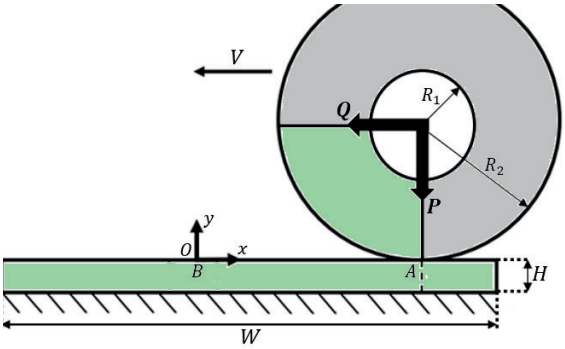


Figure 4: Schematic diagram of boundary conditions applied on the 3D transient FE model.

The roller has been exposed to a vertical force (P) and a traction force (Q) from its outer cylindrical surface. Due to the effect of the traction force, the roller is set to roll from position A toward the position B. Both the vertical force and the traction force are applied as ramped from zero to corresponding maximum value during the analysis. The relation between the vertical force and the traction force is established, as in Eq. (1). Q^* is defined as the traction coefficient that may vary between 0 and 1, according to the desired amount of traction on the roller. If traction coefficient $Q^* = 0$, it means that there is no traction, which results in a fully stick condition in contact solution. If $Q^* = 1$, it corresponds to a fully slip condition. Other than these two special cases, both the slip and stick regions are observed at the same contact patch depending on the value of the traction coefficient Q^*. In this study, four different FE analyses are performed for traction coefficients $Q^* = 0.25$, 0.50, 0.75 and 1. The analyses are performed under a constant coefficient of friction, $\eta = 0.3$. Coulomb's law of friction is followed, and the static as well as kinematic coefficients of friction are taken equal independently from the sliding velocity of the roller. Loading and friction parameters for each analysis are given in Table 3.

$$Q^* = Q/(\eta P) \quad (1)$$

Table 3: Loading, traction and friction parameters.

Analysis #	Q^* [N]	Q [N]	P [N]	η
1	0.25	112.5	1500	0.3
2	0.50	225	1500	0.3
3	0.75	337.5	1500	0.3
4	1	450	1500	0.3

Since it is a transient process, determination of the analysis end time is important. The analysis end time should be chosen wisely. If it is assigned much more than it should be, then the roller will overspeed so that the desired contact results cannot be caught. If it is assigned much less than it should be, then the stress values in the contact patch cannot reach the expected values. In this study, the analysis end time is taken as 2.5E-3 seconds, determined by trial-and-error, which corresponds to a roller speed of around 2000 rpm. Each analysis takes about 4 hours with a 28 core Intel Xeon E5 2683-v3 @2.00GHz.

III. RESULTS AND DISCUSSION

The contact analysis results corresponding to $Q^* = 0.25$, 0.50, 0.75 and 1 are depicted in Figs. 5 through 8, respectively. These figures visualize the stress values in the contact patch under the specific vertical load (P) and the horizontal traction force (Q). Normal stress in the vertical direction and the shear stress are shown in the same figure for the case of each traction coefficient. The FE and analytical results [5] are compared in each figure. Note that the results presented for analytical solution are for plane stress conditions (Check it). As expected, the distribution of slip and stick regions in the contact patch agree well with the findings of the analytical solution. The slip/stick ratio gets larger, as the traction coefficient increases. Another important feature is that the contact widths consistently possess the same values of the analytical findings.

The comparison of the shear stress profiles for analytical and FE solutions indicates a good match in the results for traction coefficients $Q^* = 0.75$ and 1. For $Q^* = 0.25$ and $Q^* = 0.50$, on the other hand, the stick regions of the shear stress profiles for analytical and FE solutions do not overlap. Nevertheless, the slip regions and the peak values in the shear stress profiles present good matches with the analytical solution in each analysis.

Normal stresses (contact pressures) remain constant for all analyses, since the same vertical loading is applied on the system. Contact pressures in the normal direction do not have any dependency on the variation of traction coefficients, as expected.

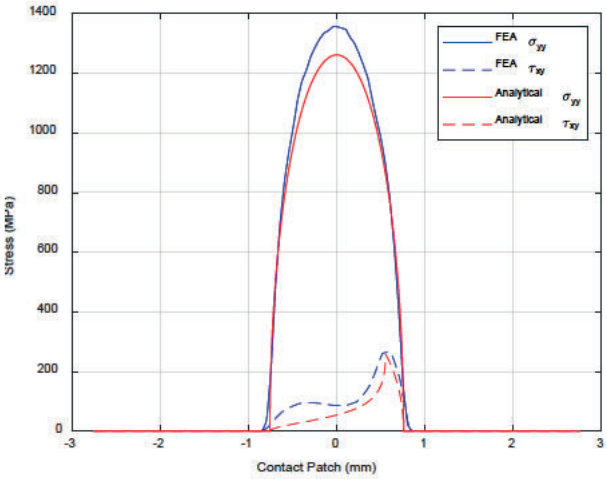


Figure 5: FEA solution vs. analytical solution for $Q^* = 0.25$.

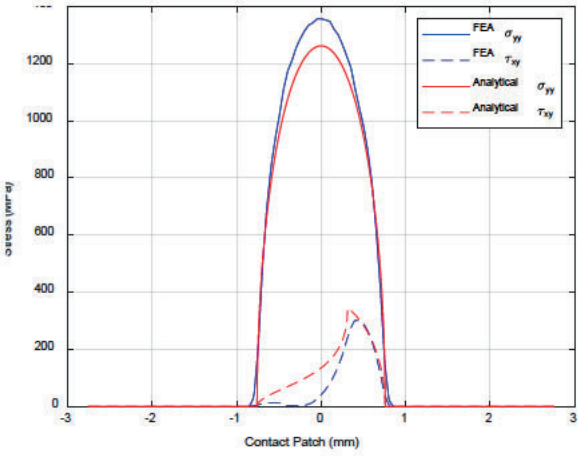


Figure 6: FEA solution vs. analytical solution for $Q^* = 0.50$.

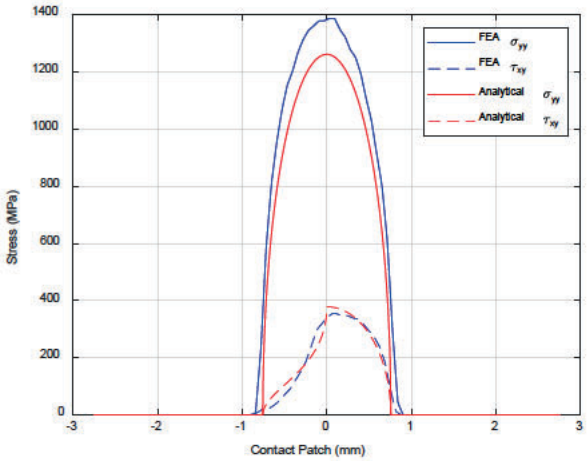


Figure 7: FEA solution vs. analytical solution for $Q^* = 0.75$.

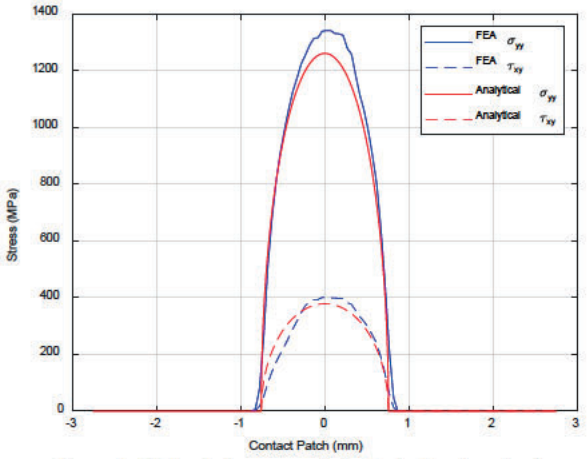


Figure 8: FEA solution vs. analytical solution for $Q^* = 1$.

Tables 4 through 7 show the error percentages for each analysis. Contact width, maximum shear stress (τ_{xy}), maximum normal stress in vertical direction (σ_{yy}) are listed and compared with the analytical results. It is seen that the general contact behavior for analytical and FE results agree well with error percentages less than 10%. The main reason behind the errors is the unwanted oscillation of the roller during its rolling movement. Since it is not a steady-state operation, the FE model cannot completely overcome these oscillations. Another reason is that no damping is applied on the system, due to the nature of the problem of interest. Discretization is also one of the crucial causes of the errors. A denser mesh can improve the precision, however it is not preferred in this study because of the burden in computational time.

Table 4: Accuracy of the FE solution for $Q^* = 0.25$.

$Q^* = 0.25$	Analytical	FEA	Error (%)
Contact width (mm)	1.51	1.61	6.6
Maximum stress, τ_{xy} (MPa)	257.8	265.5	3.0
Maximum stress, σ_{yy} (MPa)	1262	1352	7.1

Table 5: Accuracy of the FE solution for $Q^* = 0.50$.

$Q^* = 0.50$	Analytical	FEA	Error (%)
Contact width (mm)	1.51	1.62	7.3
Maximum stress, τ_{xy} (MPa)	344.5	301.1	12.6
Maximum stress, σ_{yy} (MPa)	1262	1356	7.4

Table 6: Accuracy of the FE solution for $Q^* = 0.75$.

$Q^* = 0.75$	Analytical	FEA	Error (%)
Contact width (mm)	1.51	1.66	9.9
Maximum stress, τ_{xy} (MPa)	378.5	344.5	9.0
Maximum stress, σ_{yy} (MPa)	1262	1382	9.5

Table 7: Accuracy of the FE solution for $Q^* = 1$.

$Q^* = 1$	Analytical	FEA	Error (%)
Contact width (mm)	1.51	1.62	7.3
Maximum stress, τ_{xy} (MPa)	378.5	400.7	5.9
Maximum stress, σ_{yy} (MPa)	1262	1341	6.3

IV. CONCLUSION

A finite element based methodology for the frictional rolling contact problems was presented in this paper. For the 3D transient problem, a one-step analysis method is proposed. Normal stress and contact width were determined under the effect of the applied vertical load. Various surface shear stresses distributions were obtained by applying different traction forces. Traction coefficient Q^* was applied for the various values using the following relation: $Q^* = Q/(\eta P)$ under constant coefficient of friction and constant vertical load. Depending on the traction coefficient variation, the slip/stick distribution changed as expected. These results were validated through the analytical formulation of Guler et al. (2012). This method may be used a basis for more realistic 3D solutions to the cylindrical bearing problems.

ACKNOWLEDGEMENT

The funding provided by the TUBITAK and ORS Bearings Inc. under Grant No. 5180049 is gratefully acknowledged.

REFERENCES

[1] H Hertz, Über die berührung fester elastischer körper, J Reine und Angewandte Mathematik 92 (1882) 156–171
[2] F W Carter, "On the action of a locomotive driving wheel", Proc R Soc Lond A 112 (1926) 151–157
[3] P J Vermeulen, K L Johnson, "Contact of non-spherical bodies transmitting tangential forces", J Appl Mech 31 (1964) 338–340
[4] J J Kalker, *Three-Dimensional Elastic Bodies in Rolling Contact*, Kluwer Academic Publishers, Dordrecht, 1990, pp 137–184
[5] Guler, M A , Y Alinia, and S Adibnazari "On the rolling contact problem of two elastic solids with graded coatings" International Journal of Mechanical Sciences 64 1 (2012): 62-81
[6] S Bogdanski, M Olzak, J Stupnicki, Numerical stress analysis of rail rolling contact fatigue cracks, Wear 191 (1996) 14–24
[7] Y C Chen, L Chen, "Effects of insulated rail joint on the wheel/rail contact stresses under the condition of partical slip", Wear 260 (2006) 1267–1273
[8] A Saulot, L Baillet, "Dynamic finite element simulations for understanding wheel–rail contact oscillatory states occurring under sliding conditions", J Tribol ASME 128 (2006) 761–770
[9] Zhao, X , Li, Z "The solution of frictional wheel–rail rolling contact with a 3D transient finite element model: Validation and error analysis" Wear 271 1 (2011): 444-452

[10] Zhao, X , Li, Z "A three-dimensional finite element solution of frictional wheel–rail rolling contact in elasto-plasticity " Journal of Engineering Tribology 229 1 (2014): 86-100
[11] Zhao, X , Li, Z "A solution of transient rolling contact with velocity dependent friction by the explicit finite element method " Engineering Computations 33 4 (2016): 1033-1050
[12] ANSYS, Inc , ANSYS Mechanical APDL Structural Analysis Guide, ANSYS Ver 19, 2019

Determining the Effect of Bird Parameters on Bird Strikes to Commercial Passenger Aircraft Using the Central Composite Design Method

Z. HASILCI[1] and M. BOGOCLU[2]

[1]Yildiz Technical University, Turkish Airlines Technic INC., Istanbul/Turkey, zhasilci@thy.com
[2]Yildiz Technical University, Istanbul/Turkey, bogoclu@yahoo.com

Abstract **- Today, bird strike is one of the most threatening problems to flight safety. A bird strike damage in flight can result in serious structural damage or even fatal accidents. A bird strike model requires high computational power for model preparation and nonlinear explicit analysis because of composite materials, contact definitions and other complex analysis parameters. Investigating the effects of design parameters on bird strike is a costly and time-consuming practice. The influence of various parameters such as bird velocity and impact angle has been also evaluated on a composite target in this research.**

Investigation of the effects of bird parameters on a composite target provides a clearer definition of the strength limits and energy transfer of composite materials exposed to bird strikes. Real bird strike tests are in good agreement with Ls-Dyna analysis results in this study. The unique aspect of this research is that the Central Composite Design (CCD) method, one of the Experimental Design (DOE) methods, is one of the first applications in the bird strike problem. Bird strike simulations were performed in different analysis parameters based on the Central Composite Design (CCD) method and the effects of the parameters on bird strike were found with the regression equations obtained from Minitab.

Keywords **– Bird Strike, Impact Test, Smoothed Particle Hydrodynamics (SPH), Design of Experiment (DOE), Central Composite Design (CCD).**

I. Introduction

Aircraft collisions with birds and other wildlife have become an increasing concern for aviation safety in recent years. Factors that contribute to this increasing threat are increasing populations of large birds and increased air traffic by quieter, turbofan-powered aircraft. Globally, wildlife strikes destroyed over 263 aircraft (Figure 1) between 1988 and 2018, and even killed more than 282 people [1].

The Federal Aviation Administration (FAA) extendedly compiles a database of all reported wildlife strikes to U.S. civil aircraft and to foreign carriers experiencing strikes in the USA. About 220,000 strike reports have been compiled from 2,050 USA airports and 310 foreign airports for January 1990 through September 2019 (about 16,000 strikes in 2018) [2]. Some of these strikes resulted in fatal accidents. These bird strike examples from the database, presented in chronological order, show the serious impacts that strikes by birds can have on aircraft and demonstrate the widespread and diverse nature of the problem (Table 1) [2].

Table 1: Some of the human fatalities and injuries due to bird strikes.

Date	Aircraft	People	Damage
19/11/2017	Bell 407	3 deaths	Aircraft destroyed
27/06/2017	Jordan John RV7	2 deaths	Aircraft destroyed
20/04/2016	C-172	4 deaths	Aircraft destroyed
13/02/2013	Avions Fairey Tipsy Nipper	1 deaths	Aircraft destroyed
04/01/2009	Sikorksy S-76C	8 deaths, 1 injury	Aircraft destroyed
04/03/2008	Cessna Citation I	5 deaths	Aircraft destroyed
23/10/2007	Piper 44	2 deaths	Aircraft destroyed
08/07/2003	Cessna 172	2 deaths	Wings, engine, (possibly more)
10/05/2003	MD A-4N (former military)	1 deaths	Aircraft destroyed
04/03/1998	Piper 23	2 deaths	Aircraft destroyed
15/07/1994	Cessna 172	1 deaths	Aircraft destroyed
16/05/1994	Bell BHT-47	1 deaths	Aircraft destroyed
05/06/1992	Starduster SA 300	1 deaths	Aircraft destroyed

Figure 1: Aircraft destroyed due to bird strike [3].

On 7 November 2007, a large blue heron struck the left wing leading edge of the B-737 aircraft, causing damage to the aircraft (Figure 2) [3].

Figure 2: Bird strike damage to B737 aircraft left wing leading edge [3].

Figure 3: Turkish Airlines aircraft in a flock of birds [4].

With the realization of the importance of accidents caused by bird strikes at Turkish Airlines, reporting and research on this issue has gained momentum. Figure 3 shows a situation related to the accidental diving of an aircraft in the Turkish Airlines fleet into a bird flock.

European Aviation Safety Agency (EASA) CS-25.631 and Federal Aviation Administration (FAA) FAR-25.571-e-1 (Bird Strike Damage) requires that the aeroplane must be designed to assure capability of continued safe flight and landing of the aeroplane after impact with a 4 lb bird when the velocity of the aeroplane (relative to the bird along the aeroplane's flight path) is equal to V_C (Design Cruise Speed) at sea-level or 0.85 V_C at 2438 m (8000 feet), whichever is the more critical. Compliance may be shown by analysis only when based on tests carried out on sufficiently representative structures of similar design [5].

It is common practice today to design and test aircraft components as bird proof, redesigning and retesting until the relevant component passes the tests. Bird strike tests with real birds are costly and time consuming, instead finite element analysis methods such as Ls-Dyna are used in recent years, there are many academic studies in this field [6-13]. However, so far, there has been no academic study examining the impact angle and impact velocity effects using the Central Composite Design method. The determination of the extent to which bird velocity and impact angle parameters affect outputs such as deformation, kinetic energy and internal energy with the regression equation is one of the unique aspects of this research.

II. Impact Theory

The projectile has a hydrodynamic behaviour during the impact. Soft body impact will be defined as impact in which the stresses generated substantially exceed the strength of the projectile but are well below the strength of the target material. This implies that soft body projectiles will flow upon impact while the target may see little or no plastic deformation [7].

The bird strike event is often considered as a jet of water hitting a target. It can be divided into two stages: the initial shock and the steady flow. The pressure of the initial shock (Hugoniot pressure) is given by equation (1); the pressure of the steady flow (stagnation pressure) is calculated according to Bernoulli and is given by equation (2) [6]:

Hugoniot pressure:

$$P_{shock} = \rho_0 \, v_{shock} v_{impact} \quad (1)$$

Stagnation pressure:

$$P_{stagnation} = \frac{1}{2} \rho_0 \, v^2{}_{impact} \quad (2)$$

Analytically, those two pressures are important since the Hugoniot pressure gives the maximum possible value for the impact and the stagnation pressure gives the expected reading when the flow stabilizes. It is also important to realize that the pressure is independent of the size of the projectile since the mass is not a variable in the pressure equation. So while the force and energy of a larger projectile will cause more damage, the pressure results are the same for a bird of different weight. The values of the variables needed to calculate the stagnation pressure are easily available. On the other hand, the Hugoniot pressure depends on the impact velocity and the shock velocity, which itself also depends on the impact velocity. The information required to calculate the pressures are found in Wilbeck [7].

As can be seen in Equations (1) and (2), the bird velocity is closely related to the Hugoniot pressure and stagnation pressure formed on the plate. The angle of impact is a parameter that affects the bird velocity in the normal direction of the plate. All in all, bird velocity and impact angle are two main parameters affecting the Hugonoit pressure. Therefore, this study will examine the pressure distribution over the target as a result of bird strikes at different velocities and different impact angles.

III. Smoothed Particle Hydrodynamics (SPH) Bird Model

When the studies in the literature are investigated, the most commonly used bird model is SPH (the Smooth Particles Hydrodynamics Model), because it gives closer results to experimental data than other analysis methods [9-13].

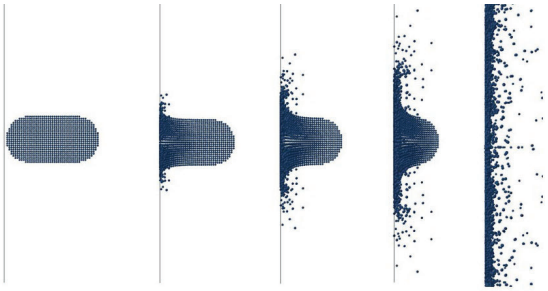


Figure 4: Deformation of the SPH bird model.

Compared to the conventional solid Lagrangian mesh, where the explicit time step decreases significantly with element deformation, the time step is constant for the SPH model. However, a sufficient particle density is required in order to

achieve accurate results, which may necessitate high memory resources and typically is a compromise between accuracy and required CPU time [8]. The stages of the bird striking modeled as SPH are shown in Figure 4. In this study, an artificial bird was modeled using 15013 SPH nodes. Since 1.8 kg bird is used, the weight of each node is 0.119896 gr.

In general, compared to the Eulerian model, the SPH method requires fewer elements, avoids the material interface problems associated with it and normally has a shorter solution time. The numerical robustness, compared to the conventional Lagrangian mesh with its mesh distortion problems, is very high [8].

Composite laminates possessing high in-plane strength and stiffness are rather sensitive to damage initiated by transverse impact loads that can cause dents or material penetration by the projectile. Depending on the impact energy determined by the projectile mass and velocity and the properties of laminate impact loading can result in considerable reduction in material strength under tension, compression, and shear.

IV. The Verifying of The Bird Strike Simulations on Composite Target

The bird strike analysis was carried out via Ls-Dyna Prepost and Ls-Dyna Manager in the current study. To accurately predict the response of an aircraft structure under impact loading, it is essential to have an accurate bird model [6]. In this reseach, the bird is modeled as a cylinder with two hemispherical ends, as shown in Figure 4. The ratio of the length to the diameter of the bird is selected to be 2:1. The weight of the bird is 1.8 kg according to EASA CS-25.631 and FAA FAR-25.571-e-1 [4, 5]. The length of the bird is 228 mm and its diameter is 114 mm [14].

When the studies are examined in the literatüre [9-14], it is seen that many studies verify the model with the real bird strike test results in Wilbeck [7].

Convergence studies were performed determining the appropriate mesh size for obtaining accurate kinetic energy, Hugoniot pressure and stagnation pressure. As the appropriate mesh size, the mesh size of the composite plate is determined as 7.7 mm in the x and y direction and 8.47 mm in the z direction.

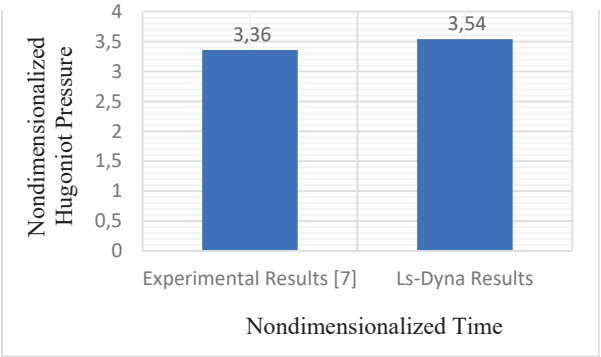


Figure 5: Comparison of real bird strike test results [7] with analysis results in terms of nondimensionalized Hugoniot pressure - nondimensionalized time.

The real bird was impacted perpendicularly at 116 m/s in the verification study as in actual test conditions. Typical pressure histories recorded at the center of impact for real bird strike are given in Figure 5 & 6. The obtained pressure values were divided by $P = {}^{1}/_{2}\,\rho v^2$ and the time was divided by $T_s = {}^{l}/_{v}$ (l: bird's length, ρ: bird's density, v: bird's velocity). The experimental [7] and analysis results of normalized Hugoniot pressure value is 3.36 and 3.54, respectively (The difference is 5.36%). The value of the normalized Hugoniot pressure is consistent with the value obtained from the experimental results [7], as clearly demonstrated in Fig. 5.

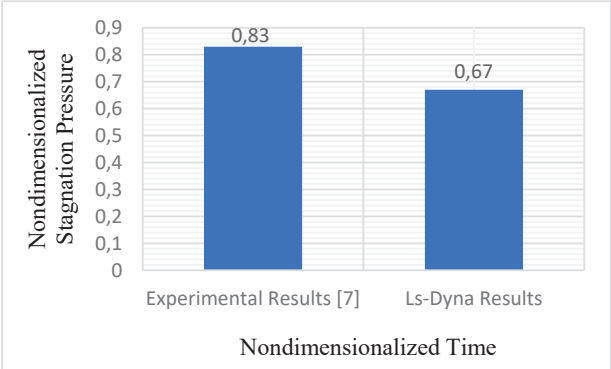


Figure 6: Comparison of real bird strike test results [7] with analysis results in terms of nondimensionalized stagnation pressure - nondimensionalized time.

The normalized stagnation pressure value obtained from this study and experimental results [7] is 0.67 and 0.83, respectively. The difference between the simulation and the experiment of bird strike is 19.28%. This difference arises from the differences between tests and experiments.

The analyzes have also been verified by comparing the kinetic energies. At the start, the analyses of bird strike were performed perpendicularly in different bird velocities. Bird's kinetic energy analysis results for bird velocity 116 m/s, 134 m/s and 152 m/s are 12100 J, 16200 J and 20800 J, respectively (Table 2). As expected, the kinetic energy of the bird increases as the velocity of bird increases. From the Table 2, it can be seen that the error percentage between the theoretical kinetic energy and Ls-Dyna result is about 0.086%, 0.244% and 0.031% respectively. This result shows that the bird strike model has a high predictive capability, and the bird strike model can replace the real bird strike tests.

Table 2: Comparison of the simulation with the theoretical kinetic energy (V_b: bird velocity [m/s]).

V_b [m/s]	Simulation [J]	Theoretic [J]	Error (%)
116	12100	12110,40	0.086
134	16200	16160,40	0.244
152	20800	20793,60	0.031

For applying the boundary conditions of the target during analysis, the edges of the target is fixed. The properties of T700/M21 carbon fiber material used in the analyzes are given in Table 3.

In this study, the composite target is 36 ply and its stacking sequence is $[0_2/45_2/90_2/-45_2/0_2/-45_2/0_2/45_2/0_2]_s$. Such a quasi-isotropic ply orientation simulates the properties of an isotropic material. It is therefore advisable to build a symmetrical and balanced laminate to minimize any residual thermal stresses induced during resin curing.

Table 3: The properties of T700/M21 carbon fiber epoxy prepreg material [13].

Mechanical Constants (Symbols)	Values
Density (ρ)	1600 kg/m^3
Young's Modulus (E_A)	130 e+9 Pa
Young's Modulus (E_B)	7.7 e+9 Pa
Young's Modulus (E_C)	7.7 e+9 Pa
Poisson's Ratio (ν_{BA})	0.0195
Shear Modulus (G_{AB})	4.8 e+9 Pa
Shear Modulus (G_{BC})	4.8 e+9 Pa
Shear Modulus (G_{CA})	4.8 e+9 Pa
Ultimate Tensile Stress (X_T)	2080 e+6 Pa
Ultimate Tensile Stress (Y_T)	60 e+6 Pa
Ultimate Compressive Stress (X_C)	1100 e+6 Pa
Ultimate Compressive Stress (Y_C)	180 e+6 Pa
Ultimate Shear Stress (S_C)	110 e+6 Pa
Failure Criterion ($CRIT$) / Tsai-Wu	55

V. Response Surface Design

In studies in the literature [15-21], typically, response surface methodology (RSM) is used to optimize the process parameters in casting, welding and machinability studies of composite materials. In this study, different from the studies in the literature, the significant variables which are bird velocity and impact angle are examined using response surface methodology (RSM) based on central composite design (CCD). Response surface design is two types: Central Composite designs and Box-Behnken designs.

A. Central Composite Design Method

Compared to other DOE methods, the Central Composite Design method is the most suitable method for this problem. Central Composite Design method minimises the number of experiments for a specific number of factors and its levels. Experiments are conducted as per the experimental design matrix and the output responses are recorded for analysis. The process parameters which significantly influence the response are identified using analysis of variance. Regression equations are used to predict the response for the given process parameters and its levels. Response surface plots are used to understand the effect of process parameters on response. Finally, the process parameters are optimised using desirability approach of response surface methodology and confirmed by conducting confirmation tests [15]. Central composite designs are a factorial or fractional factorial design with center points, augmented with a group of axial points that allow the curvature to be predicted. Central composite designs are especially useful in sequential experiments because that can be build on previous factorial experiments by adding axial and center points [22]. After conducting bird strike analysis in two different parameters in Ls-Dyna, the Central Composite Design method was applied.

Table 4: Z displacement, penetration and pressure results in Ls-Dyna.

Input		Output		
Bird Velocity [m/s]	Impact Angle	z disp. [mm]	Penetration	Pressure in the Center of Plate [Mpa]
116	0	180,00	yes	314
116	30	33,30	no	298
116	45	12,82	no	291
116	60	5,09	no	298
134	0	223,00	no	319
134	30	153,05	yes	318
134	45	15,69	no	327
134	60	5,60	no	186
152	0	254,70	yes	207
152	30	239,93	yes	315
152	45	100,76	yes	291
152	60	6,27	no	299
108,544	30	25,27	no	305
134	-12,4264	242,78	yes	310
134	72,4264	1,46	no	155
159,456	30	281,59	yes	316

The Central Composite Design method was applied after 12 different bird strike analysis was done in Ls-Dyna, with 2 different design parameters, 3 levels for bird speed (116 m/s, 134 m/s and 152 m/s) and 4 levels (0^0, 30^0, 45^0 and 60^0) for impact angle. Twelve different test were prepared, these analyzes were analyzed in Minitab with the Central Composite Design method. As a result, four different analyzes were obtained from Minitab. The total number of analyzes which is performed is sixteen in different impact angles and bird velocities. These sixteen different analyzes were conducted at Ls-Dyna. Analysis results are presented in Table 4 and Table 5. The z-direction deformation of the plate and the pressure values in the center of the plate depending on the bird velocity and the angle of impact are shown in Table 4. Table 5 presents the kinetic energy of plate, the plate of internal energy and the bird of internal energy values from the bird strike analysis results.

Table 5: Kinetic energy and intertial energy in bird strike analysis.

Input		Output		
Bird Velocity [m/s]	Impact Angle	The Kinetic Energy of Plate [J]	The Plate of Internal Energy [J]	The Bird of Internal Energy [J]
116	0	1310	7890	444
116	30	708	2680	370

116	45	471	1380	276
116	60	220	668	147
134	0	2520	10700	509
134	30	1910	10200	631
134	45	701	2290	354
134	60	346	968	196
152	0	3360	12000	836
152	30	3100	10100	476
152	45	872	8300	573
152	60	504	1400	256
108,544	30	594	1870	326
134	-12,4264	2360	10800	455
134	72,4264	112	312	80,5
159,456	30	3910	16000	699

VI. Result And Discussion

A. Internal Energy of Plate

The accuracy of the bird strike model in this study by using three statistical parameters. These parameters are the coefficient of the multiple determination R-sq, root mean squared error (RMSE) and then adjusted coefficient of determination R-sq (adj), respectively. The R-sq and R-sq (adj) values of around 80% indicate that the model fits well with the bird strike model in Ls-Dyna (Table 6).

Table 6: Plate's internal energy DOE model summary.

S	R-sq	R-sq(adj)	R-sq(pred)
2361,75	85,87%	78,80%	52,54%

Regression Equations (4) and (5) are obtained from Minitab (V_b: Bird Velocity [m/s], α: Impact Angle [deg]). The internal energy of the plate is min 312 J, max 16000 J from Equation (4).

$$Internal\ Energy\ Plate\ [J] = -2702 - 32\ V_b + 113\ \alpha + 0{,}96\ {V_b}^2 - 1{,}33\ \alpha^2 - 1{,}33\ V_b * \alpha \quad (4)$$

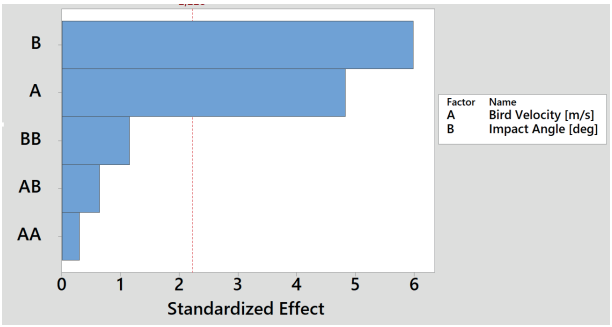


Figure 7: Pareto chart of the standardized effects (response is the plate of kinetic energy and internal energy for $\alpha = 0.05$).

Information about the effect of each input parameter on bird strike analysis results is provided in the Figure 7. According to Figure 7, at 95% confidence interval, it is the angle of impact that most affects the kinetic energy and internal energy results.

Kinetic energy, internal energy and z displacement are output in bird strike analysis. Figure 8 shows internal energy predicted from the analysis results as the bird velocity and the angle of impact change.

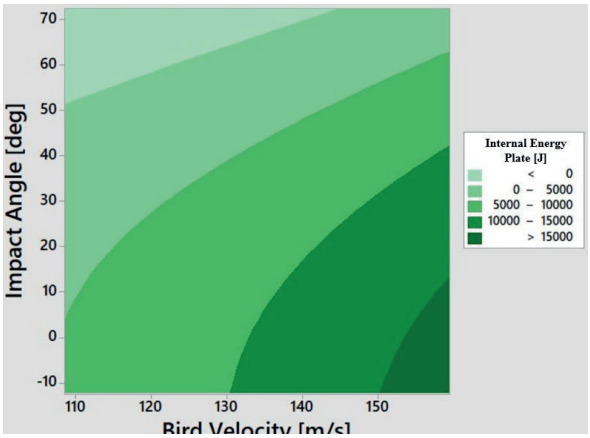


Figure 8: Contour plot of internal energy.

There are two factors: bird velocity and impact angle. The range of bird velocity in the analysis is 108,544 m/s and 159,456 m/s. The range of the impact angle is -12.4264^0 and 72.4264^0 degrees.

As we consider the values in the data set, while the bird velocity is 159.456 m/s and the angle of impact is -12.4264^0, the internal energy of the plate is at the maximum value and is 17607.9 J.

B. Kinetic Energy of Plate

The bird of theoretical kinetic energy is calculated $0.5 * m * {V_b}^2$. The min. kinetic energy value formed on the plate is 112 J, and the max. value is 3910 J from Equation (5).

$$Kinetic\ Energy\ Plate\ [J] = 2239 - 77\ V_b + 117{,}0\ \alpha + 0{,}578\ {V_b}^2 - 0{,}271\ \alpha^2 - 0{,}999\ V_b * \alpha \quad (5)$$

The R-sq and R-sq (adj) values of around 80% indicate that the model fits well with the bird strike model in Ls-Dyna (Table 7). It can be concluded that the numerical results of kinetic energy also perfectly matched with the theoretical bird strike analysis results (Table 7).

Table 7: Plate's kinetic energy DOE model summary.

S	R-sq	R-sq(adj)	R-sq(pred)
511,532	88,71%	83,07%	67,92%

Figure 9 presents kinetic energy predicted from the analysis results depending on bird velocity and impact angle. According to this figure, the most damage occurs on the composite material at high bird velocity and when the bird is impacted perpendicularly.

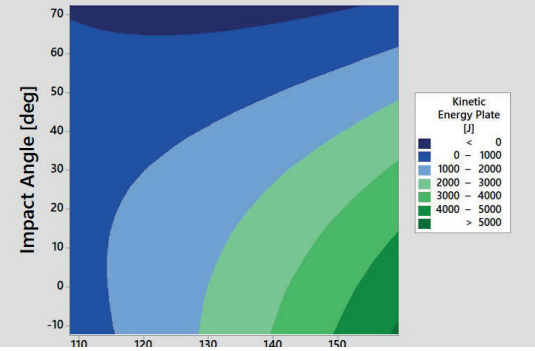


Figure 9: Contour plot of kinetic energy.

VII. Conclusion

In this study, the damage caused by the bird strike problem, which is one of the most risky issues for aviation, on composite materials and energy transfer issues were examined.

A bird strike model requires high computational power for model preparation and nonlinear explicit analysis because of composite materials, contact definitions and other complex analysis parameters.

Investigation of the effects of bird parameters on a composite target provides a clearer definition of the strength limits and energy transfer of composite materials exposed to bird strikes.

Firstly, the numerical results were compared with theoretical and experimental studies in the literature and verified as a result of iterations. Secondly, the Ls-Dyna analysis was expanded by considering the design variables. The bird velocity and impact angle were considered as design variables.

Central Composite Design (CCD) method was very effective in predicting analysis results of the composite target which was subjected to bird strike. Investigating the effects of design parameters on bird strike is a costly and time consuming practice. In this study, the result values were estimated using the Central Composite Design method in order to get a faster response, and with the regression equations obtained as a result of the analysis, the effect of which parameter on the bird strike event was found. The good correlation between the experimental data and the Ls-Dyna simulations confirms that the properties used for the bird in the numerical simulations are appropriate. In future studies, other analysis results can be estimated using the same methodology. Thus, analysis results can be predicted with a certain accuracy in a shorter time via scholastic methods.

Acknowledgment

This research was supported by the Yildiz Technical University Scientific Research Projects Coordination Department (Project number: FDK-2020-3851). The authors would like to thank the Yildiz Technical University Scientific Research Projects Coordination Department for their financial support.

References

[1] The Federal Aviation Administration (FAA), Wildlife Strikes to Civil Aircraft in the United States, 1990–2018

[2] FAA Wildlife Strike Database, U S Department of Agriculture Animal and Plant Health Inspection Service, Some Significant Wildlife Strikes to Civil Aircraft in the United States, 1990 – 2019, USA

[3] R A DOLBEER and S E WRIGHT, FAA Wildlife Strikes to Civil Aircraft in the United States, University of Nebraska - Lincoln, 1990 – 2007

[4] News: [Online], Available: https://artigercek com/haberler/thy-ucagi-kus-surusune-carpti

[5] European Aviation Safety Agency (EASA), Certification Specifications and Acceptable Means of Compliance for Large Aeroplanes, EASA / CS-25 631, Amendment 18, 2016

[6] Lavoie, M A , Gakwaya, A , Ensan, M N and Zimcik, D G , "Review of existing numerical methods and validation procedure available for bird strike modelling", ICCES, vol 2, no 4, pp 111-118, 2007

[7] J S Wilbeck, Impact behavior of low strength projectiles, U S Air Force, (1978), Report AFML-TR-77-134

[8] Sebastian Heimbs, "Computational methods for bird strike simulations: A review", Computers and Structures 89, pp 2093–2112, 2011

[9] A F Johnson, M Holzapfel, Modelling soft body impact on composite structures, Compos Struct 61, pp 103–113, 2003

[10] M A McCarthy, J R Xiao, C T McCarthy, A Kamoulakos, J Ramos, J P Gallard, V Melito, Modelling of Bird Strike on an Aircraft Wing Leading Edge Made from Fibre Metal Laminates – Part 2: Modelling of Impact with SPH Bird Model, Applied Composite Materials, 11, pp 317-340, 2004

[11] M Nejad Ensan, D G Zimcik, M, Lahoubi, D Andrieu, Soft Body Impact Simulation on Composite Structures, Transactions Canadian Society for Mechanical Engineering, 2008

[12] M A Lavoie, A Gakwaya, M N Ensan, D G Zimcik, D Nandlall, Bird's substitute tests results and evaluation of available numerical methods, Int J Impact Eng 36, pp 1276–1287, 2009

[13] Y Zhou, Y Sun and T Huang, "Bird-strike resistance of composite laminates with different materials", The Journal of Materials, 13, 129; December 2020

[14] I Smojver, D Ivančević, "Numerical simulation of bird strike damage prediction in airplane flap structure", Composite Structures, 92, pp 2016–2026, 2010

[15] S J S Chelladurai, M K , A P Ray et al , "Optimization of process parameters using response surface methodology: A review", Materials Today: Proceedings, June 2020

[16] H Joardar, N S Das, G Sutradhar, S Singh, "Application of response surface methodology for determining cutting force model in turning of LM6/SiC_P metal matrix composite", The Journal of Measurement, 47, pp 452–464, September 2014

[17] S Gopalakannan, T Senthilvelan, "Application of response surface method on machining of Al–SiC nano-composites", The Journal of Measurement 46, pp 2705–2715, April 2013

[18] R AROKIADASS, K PALANIRADJA, N ALAGUMOORTHI, "Prediction and optimization of end milling process parameters of cast aluminium based MMC", Trans Nonferrous Met Soc China 22, pp 1568-1574, February 2012

[19] H Oktem, T Erzurumlu, H Kurtaran, "Application of response surface methodology in the optimization of cutting conditions for surface roughness", Journal of Materials Processing Technology 170, pp 11–16, April 2005

[20] S J S Chelladurai, R Selvarajan, T P Ravichandran, S K Ravi, and S R C Petchimuthu, "Optimisation of dry sliding wear parameters of squeeze cast AA336 aluminium alloy: copper-coated steel wire-reinforced composites by response surface methodology", International Journal of Metalcasting, September 2018

[21] K Palanikumar , N Muthukrishnan & K S Hariprasad, "Surface roughness parameters optimization in machining A356/SiC/20p metal matrix composites by PCD tool using response surface methodology and desirability function", Machining Science and Technology: An International Journal, 12:4, pp 529-545, December 2008

[22] Minitab 18 Support Training Document [Online], Available: https://support minitab com/en-us/minitab/18/help-and-how-to/modeling-statistics/doe/supporting-topics/response-surface-designs/response-surface-central-composite-and-box-behnken-designs/

Experimental study on penetration and perforation behavior of aluminum honeycomb sandwich structures using energy profiling method

M.E. ÇETİN[1] and A. MERAM[2]

[1] Necmettin Erbakan University, Konya/Turkey, mecetin@erbakan.edu.tr
[2]KTO Karatay University, Konya/Turkey, ahmet meram@karatay.edu.tr

Abstract **- The honeycomb sandwich panels are being used in aerospace, automotive, railway and marine industries due to their high flexural strength, lightweight and energy absorbing capacity. This study aimed to investigate the penetration properties of the sandwich panels using low velocity impact tests conducted according to ASTM D7136 standard. The honeycomb core and skin of sandwich panels were fabricated from Aluminum alloy series 3005 H19 and 5754, respectively. The dimension of sandwich panel specimen is 100x100x20 mm and clamped from their 75x75 mm area. The projectile has 24 mm semispherical tip and mass of 20 kg. Sandwich panels was impacted at 25, 50, 75, 100, 125, 150 and 175 J energy levels. Force-deflection data obtained from the tests and used for evaluating energy absorption capacity of sandwich panels. Energy profiling method was conducted to determine the penetration and perforation thresholds of sandwich panels under low velocity impact. Front and rear side view near the impact area in tests were investigated. It is found that 150 J energy level is the perforation threshold of sandwich panels. The study results show how the energy levels affect the penetration and perforation and provide useful knowledge for its applications in engineering.**

Keywords **- Honeycomb core, Energy profile method, Sandwich panel;**

I. Introduction

Recently, honeycomb sandwich panel have been widely employed in aerospace, marine, railway, automotive, and industries due to light weight and high stiffness [1]. These structures consist of a honeycomb core with high stiffness and two skin in top and bottom. During application, due to accidents such as bird strike, tools dropping, indentation and perforation may happen in these structures. Due to this fact, many researchers have investigated the penetration and perforation behavior of the sandwich panels under quasi-static and dynamic loadings. For instance, Paik et al. [2] conducted low-velocity impact tests to study the effects of honeycomb foil thickness on energy absorption and indentation of the sandwich panels. Crupi et al. [3] investigated the low velocity impact and quasi-static indentation behavior of the sandwich structures with foam cores. Sun et al. [4] researched the quasi-static indentation/perforation responses of honeycomb sandwich panels by considering the failure modes and energy absorption of panels in various loading stages and structural parameters. Rahimijonoush and Bayat [5] numerically and experimentally analyzed the ballistic limit velocity of aluminum honeycomb sandwich panels with titanium facesheets. Meram et al. [6] experimentally studied the effect of adhesive shear strength on failure mechanism and impact behavior of sandwich panels under different initial impact energies. Most of the previous studies [7-10] have been mainly considered the effect of the structural parameters of the sandwich panels or loading condition on the impact or indentation/perforation behavior of the structure. This study experimentally investigates the penetration and perforation thresholds of sandwich panels under low velocity impact using energy profiling method.

II. Material And methods

The sandwich panels used in low velocity impact tests manufactured by Altıgen Space Aviation Ship Construction Inc., Konya, Turkey. Sandwich panels consisted of aluminum 3005-H19 foil sheet with 0.05 mm thickness, 10.4 mm cell size as the honeycomb core; and aluminum 5754 as the top and bottom face sheets with 1 mm thickness. Honeycomb core and facesheets were bonded with a two component KLB-75 polyurethane adhesive. For impact tests, specimens were prepared with a 100 x 100 x 20 mm square dimension. The sandwich panel specimen was clamped on a fixture along a square circumference having a 75 mm side.

The low velocity impact tests were performed using a drop weight test system (Figure 1). The dropping mass was 20 kg. A semi-spherical projectile nose of 24 mm in diameter was used. Data acquisition and processing system provided by a National Instruments Signal Express data acquisition software. The load values were measured by using a 22.6 kN quartz force ring (PCB Quartz ICP force sensor 200B04). The dropping mass was released from 0.25, 0.375, 0.5, 0.675 and 0.75 m heights to produce 50, 75, 100, 125, 150 and 175 J kinetic energy in impact instance. The anti-rebounding mechanism was triggered following the initial impact to prevent the secondary impact. The load-time history for each specimen was recorded as shown in Figure 2. The numerical integration method and Newton's second low of motion were used to extract the load-deflection data from load-time data. These values are utilized to evaluate

the penetration and perforation behavior of honeycomb sandwich panels.

Figure 1: Drop weight test system

III. Results and Discussion

Low velocity impact tests were applied at 50 J, 75 J, 100 J, 125 J, 150 J and 175 J energy levels in order to obtain the penetration and perforation behavior of sandwich panels. Load-time histories acquired from impact tests was shown in Figure 2. Impact duration and maximum impact force can be depicted from this graph. It is obvious that maximum impact force in specimen 125 J is higher than the specimens which are impacted at 50, 75, 100, 150 and 175 J energy levels. Increasing the impact energy caused increase in maximum impact force until 125 J energy level. But increase in energy to 150 J and 175 J values results with decrease in maximum impact force. This situation can be explained by the higher velocity causes higher deformation takes place in more contact time. But in 150 J and 175 J energy levels perforation occurred and impact energy is used for bending and penetration resistance of the bottom face sheet. Therefore, maximum impact force decreased. Occurrence of a damage gives rise to the oscillations of the curves and rises with impact energy.

Damage propagation during low velocity impacts tests can be determined using characteristic of load-deflection graphs. Load-deflection curves of specimens for 50 to 175 J impact energies were shown in Figure 3. Though, curves of 50 to 125 J impact specimens had similar characteristics, only curve of 150 and 175 J impact specimen shows unusual behavior.

In both curves, the load increased linearly until initial peak load. Semi linear part of the load-deflection curve demonstrates the bending rigidity and specimen performs resistance against impact load. As the impact energy increases bending behavior was almost same for all energy levels. In this area face-sheet damaged and then load dropped with oscillations. In this manner crushing and buckling took place in the honeycomb core. Then as impact energy increases global rigidity increases and load starts increasing again.

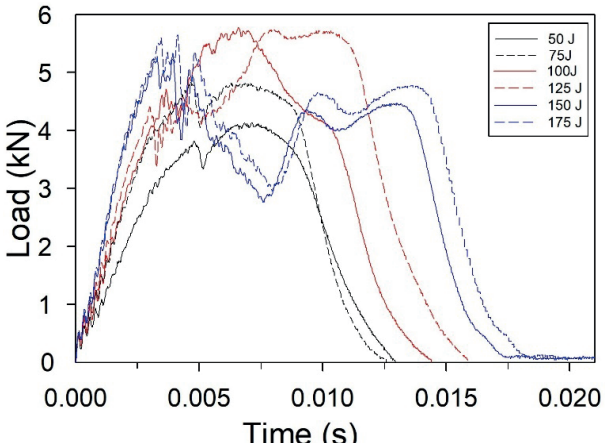


Figure 2: Load-time histories of low velocity impact tests

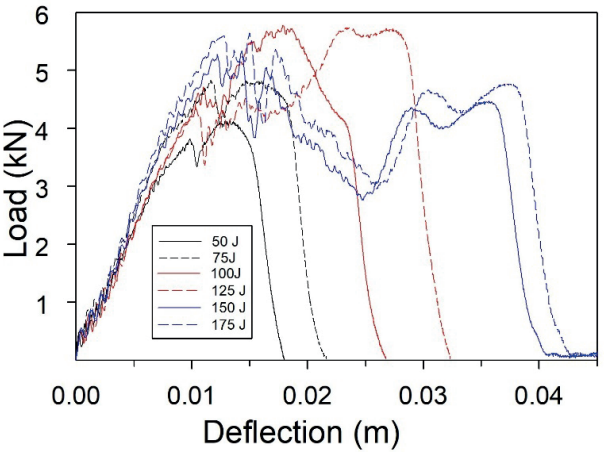


Figure 3: Load-deflection curves of low velocity impact tests

Energy profile diagram of aluminum honeycomb sandwich panels is shown in Figure 4. Equal energy line is a diagonal line and represents similarity between absorbed and impact energies. This diagram indicates the relation between the impact energy and absorbed energy [11-12]. This approach is useful for comparison of the absorbed and impact energies and also determining the thresholds of the penetration and perforation [11-12]. The energy introduced to a panel from the projectile is defined as impact energy and the energy absorbed by the panel at the final of the impact event is defined as absorbed energy [11-12]. Total absorbed energy is calculated by integrating the load-deflection curve respect to the deflection. The absorbed energies for impact tests with 50, 75,

100, 125 and 150 J impact energy levels are 46.44, 68.49, 94.58, 120.53 and 149,1 J, respectively. The comparison of the absorbed energy values of specimens show that as impact energy increases absorbed energy values increases. After the perforation, by increasing the impact energy, the absorbed energy remains approximately constant.

When the sandwich panel perforated by the projectile at the end part of load-deflection graph descending section a friction section between the projectile and the bottom sheet induces [11-12]. In addition, this part of load-deflection curve is identified as post-perforation section [11-12]. For this study, at 150 and 175 J energy levels this phenomenon happened and it is obvious from load-deflection and energy profile diagram curves. It is understood from the curves that impact energy value up to 125 J, the absorbed energy value is lower than impact energy. But in 150 and 175 J energy levels absorbed energies remain constant. The variation between absorbed energy and impact energy retains in the projectile and rebounds the impactor from the specimen final of impact event. As a result, energy absorption capability of sandwich panels seems to be well until 125 J energy level without perforation. After 150 J impact energy the perforation is occurred in the specimens.

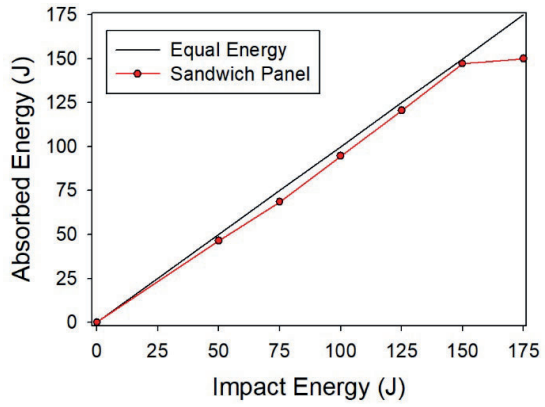


Figure 4: Energy profile diagram of sandwich panels.

Figure 5 and Figure 6 shows damaged specimen after 150 J energy level drop test. Front and rear side images of damaged specimens shows that maximum stroke and damage is on the specimen 150 J. Tearing in the top sheet and local buckling and tearing in the bottom sheet can be seen from Figure 6. In addition crushing and buckling in the honeycomb core happened.

Figure 5: Front image of sandwich panel impacted at 150 J energy level.

Figure 6: Rear image of sandwich panel impacted at 150 J energy level.

IV. Conclusion

This study investigated the penetration/perforation properties of the Aluminum honeycomb sandwich panels using low velocity impact tests conducted according to ASTM D7136 standard. The low-velocity impact tests were performed at 25, 50, 75, 100, 125, 150 and 175 J energy levels. To acquire the load-time histories, impact force sensor and data acquisition system were used. Numerical integration method and Newton's second low of motion were used to extract the load-deflection data. Energy profiling method was conducted to determine the penetration

and perforation thresholds of sandwich panels under low velocity impact. Front and rear side view near the impact area in tests were investigated. It was found that 150 J energy level was the perforation threshold of sandwich panels and after this level any impact energy results with perforation. The study results showed that the penetration /perforation behavior of honeycomb sandwich panels are highly sensitive to the impactor energy level.

ACKNOWLEDGMENT

The authors would like to thank Altıgen Space Aviation Ship Construction Inc., Konya, for their support in kindly supplying the honeycomb sandwich panels.

REFERENCES

[1] A P Meran, T Toprak, and A Muğan, Numerical and Experimental Study of Crashworthiness Parameters of Honeycomb Structures, Thin Wall Struct , 2014, 78, p 87–94

[2] Paik JK, Thayamballi AK, Kim GS The strength characteristics of aluminum honeycomb sandwich panels Thin-walled Struct 1999;35(3):205–31

[3] Crupi V, Epasto G, Guglielmino E Comparison of aluminium sandwiches for lightweight ship structures: Honeycomb vs foam Marine Struct 2013;30:74–96

[4] G Y Sun, D D Chen, X T Huo, G Zheng, Q Li, Experimental and numerical studies on indentation and perforation characteristics of honeycomb sandwich panels, Compos Struct 184 (2018) 110–124

[5] Rahimijonoush, A , & Bayat, M (2020) Experimental and numerical studies on the ballistic impact response of titanium sandwich panels with different facesheets thickness ratios Thin-Walled Structures, 157, 107079

[6] Meram, A , Çetin, M E Experimental Investigation on the Effects of Core/Facing Interface Performance on the Low-Velocity Impact Behavior of Honeycomb Sandwich Panels J of Materi Eng and Perform (2020) https://doi org/10 1007/s11665-020-05181-2

[7] Ivañez, I , Sánchez-Saez, S , Garcia-Castillo, S K , Barbero, E , Amaro, A M , & Reis, P N (2020) Impact response of repaired sandwich structures Polymer Composites

[8] Ebrahimi H, Ghosh R, Mahdi E, et al Honeycomb sandwich panels subjected to combined shock and projectile impact Int J Impact Eng 2016; 95: 1-11

[9] Yellur MR, Seidlitz H, Kuke F, et al Low velocity impact study on press formed thermoplastic honeycomb sandwich panels Compose Struct 2019; 225: 111061

[10] Ozdemir O, Karakuzu R and Al-Shamary AKJ Core-thickness effect on the impact response of sandwich composites with poly vinyl chloride and poly ethylene terephthalate foam cores J Compos Mater 2015; 4911: 1315-1329

[11] Sayer, M , Bektaş, N B , & Sayman, O (2010) An experimental investigation on the impact behavior of hybrid composite plates Composite Structures, 92(5), 1256-1262

[12] Liu, D , Raju, B B , & Dang, X (2000) Impact perforation resistance of laminated and assembled composite plates International Journal of Impact Engineering, 24(6-7), 733-746

The Effect of Leading-Edge Tubercles on Flow Around a Semi-Span NACA 0020 Airfoil

Aleyna ÇOLAK[1], Mehmet SEYHAN[1] and Mustafa SARIOĞLU[1]

[1] Karadeniz Technical University, Trabzon/Turkey, alyncolak@gmail.com
[1]Karadeniz Technical University, Trabzon/Turkey, mehmetseyhan@ktu.edu.tr
[1]Karadeniz Technical University, Trabzon/Turkey, sarioglu@ktu.edu.tr

Abstract **- In this study, the effect of the sinusoidal leading edge on the semi-span NACA 0020 airfoil models having different wavelength have been investigated experimentally in order to search aerodynamic characteristics. The experiments were carried out by measuring aerodynamic forces in the suction type wind tunnel for Reynolds number of $1.0x10^5$. Semi-span NACA 0020 airfoil models, having constant amplitude (0.025c) and two-wavelength value (0.25c, 0.5c) are used. Three different leading-edge tubercles and one baseline semi-span NACA 0020 airfoil are used in this study. Modified models produce higher lift coefficient than baseline model.**

Keywords **- NACA 0020 airfoil, leading-edge, tubercle, lift coefficient, drag coefficient.**

I. Introduction

Flow control has great importance in terms of delay flow separation, reduce noise and decrease drag coefficient. Earlier studies show that the leading-edge tubercles have been considered passive control device**.** Fish and Battle [1], suggested that the flipper which has tubercle structure can be used for flow control and showed the cross-sectional area of the humpback whale was similar to the NACA 63_4-021 airfoil. The maneuverability of humpback whales was attributed to unique structure of its pectoral flipper by the researcher. The humpback whale (Megaptera novaeangliae) is shown in Fig. 1.

Miklosovic et al. [2], that are inspired by the flipper of the humpback whale, investigated the aerodynamic effects of NACA 0020 airfoil model with sinusoidal geometry on the leading edge in the range of Reynolds number between $5.05 \times 10^5 – 5.20 \times 10^5$. Their results showed that the presence of leading-edge tubercles delays the stall angle about 40%. Miklosovic et al. [3], also used full-span and semi-span models based on the NACA 0020 airfoil profile and desired to determine the reason of airfoil performance improvement. Their results indicated that 3D effects are more prominent due to generation of vortices. Johari et al. [5] measured the lift, drag force and pitching moment on the NACA 63_4-021 airfoil at Re= 1.83×10^5 by using 6 sinusoidal models with 3 different amplitudes (0.025c, 0.05c and 0.12c) and 2 different wavelengths (0.25 and 0.50c).

Figure 1: Humpback whale (Megaptera novaeangliae) Indian Ocean, Wild Coast, South Africa, Ullstein Bild [retrieved 23 October 2020][4]

They observed an increase in lift coefficient after the stall angle and an increase in the drag coefficient before the stall angle compared to the baseline airfoil. They also performed the tuft flow visualization method to investigate the separation characteristics of models and observed that separation occurs between tubercles in modified models.

Hansen et al. [6] tried to demonstrate the effects of different airfoil profile over the aerodynamic characteristics at Re = $1.2x10^5$. When all models are examined, for NACA 65-021 airfoil tubercle structure shows much better performance when compared to NACA 0021 airfoil. For both airfoils, smaller tubercle amplitude contributes a higher maximum lift coefficient, but larger amplitude shows better performance after stall angle. Additionally, smaller wavelength improves the post-stall characteristics. According to the results, they suggested that tubercles act in a similar way as vortex generators. Post et al. [7], investigated the six different airfoil having leading-edge tubercle in the range of Re between $1.0x10^5$ - $5.0x10^5$ and performed oil flow visualization to understand the flow phenomena between the models. Results show that with leading-edge tubercle airfoil models enable to delay stall and produce higher lift about 25% at post stall region.

As shown in Figure 1, Humpback whale flipper geometry has not uniform distribution of tubercles. Therefore, in this study, models are designed to find the optimum values of the amplitude and wavelength by considering the geometry of the humpback whale flipper. The purpose of this study is to

investigate the effect of leading-edge having different wavelength on NACA 0020 airfoil.

II. METHOD

The experiments were carried out in an open circuit wind tunnel having a cross-section of 570mm x 570mm and a length of 1200 mm in the Mechanical Engineering Department of Karadeniz Technical University. Turbulence intensity of wind tunnel was obtained below 0.5% for velocity values above 3 m/s. Free stream velocity between 0 m / s and 50 m / s in test section is precisely adjusted with frequency inverter.

The experimental setup consists of endplate, connection rod, protection pipe, test model, load cell and rotary unite. Figure 2 indicates a schematic diagram of the measurement system. The gap between the endplate and airfoil model are set 1-2 mm on the bottom of the wind tunnel. Force measurements were carried out by using a 6-axis ATI Gamma DAQ F / T load cell. Measurements were taken as 20000 data at a sampling frequency of 1000 Hz. Reynolds number based on the mean chord length and experiments were carried out at $Re = 1.0x10^5$ in $0° \leq \alpha \leq 30°$. One baseline and three different NACA 0020 airfoil having leading-edge tubercles are manufactured from 3D printer. All tubercles models scaled to mean aerodynamic chord c=103 mm and half-span s=457 mm. The idealized humpback whale geometry equations are taken from the study of Miklosovic et al. [3].

Baseline model has a smooth leading edge. J1 and J2 models have a sinusoidal leading edge where the tubercles have constant amplitude. J1 has a wavelength λ_1=0.5c and J2 has λ_1=0.25c. J3 has the same amplitude and two wavelengths λ_1=0.5c, λ_2=0.25c.

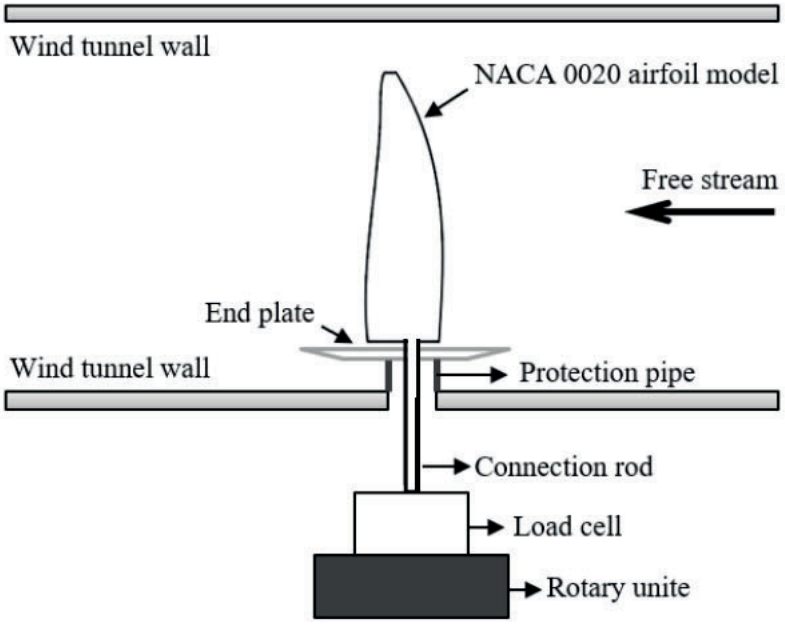


Figure 2: Schematic view of the force measurement system

$$x_{LE} = a_1 \cos\left(\frac{2\pi y}{\lambda_1}\right) \quad (1)$$

Equation (1) is used to obtain leading-edge tubercle geometry for J1 and J2.

$$x_{LE} = a_1 \cos\left(\frac{2\pi y}{\lambda_1}\right) + a_2 \cos\left(\frac{2\pi y}{\lambda_2}\right) \quad (2)$$

The leading edge geometry of the J3 model is created by using Equation (2) and used the same amplitude and two wavelengths λ_1=0.5c, λ_2=0.25c. Figure 3 shows wavy geometry of J3 obtained by combining two wave geometry.

Many studies have been carried out by changing the wavelength and amplitude to achieve maximum lift while keeping the drag minimum. While determining important parameters such as amplitude and wavelength, the study of Johari et al. [5] has been taken into account. Front view representation of tubercle models is given in Figure 4.

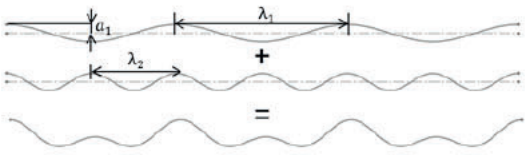


Figure 3: Schematic view of leading-edge geometry

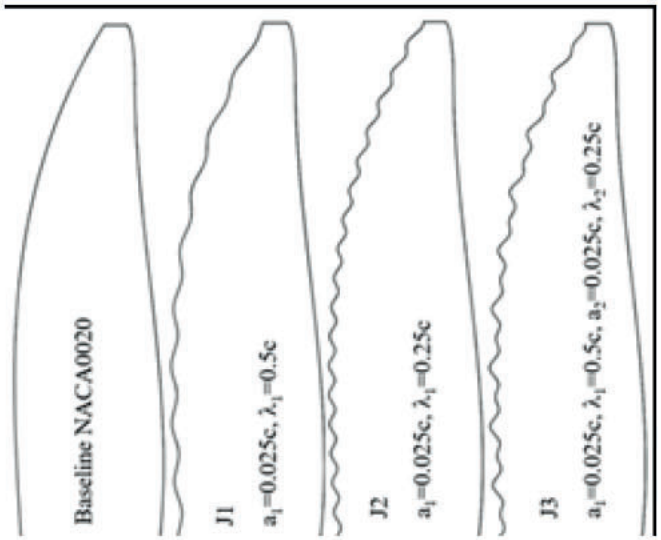


Figure 4: Front view of NACA0020 models

III. RESULTS

The aerodynamics parameters such as lift and drag coefficients of the semi-span airfoils measured and calculated at 0 deg ≤ α ≤ 30 deg. A comparison of the lift coefficient for baseline and three modified models are shown in Figure 5. It is seen from the figure the baseline and J1, J2, J3 models have similar characteristics for angles of attack below 9 deg. Before the stall angle, the effect of producing 3D flow structure with leading-edge tubercle for J1, J2 and J3 have not prominent therefore they indicate a similar trend with the baseline model. For the baseline model, stall occurring at 10 deg where the lift coefficient reaches the maximum value. Maximum lift coefficient for J1 model increased by 6 % compared to the baseline. J1 shows the same increment tendency with baseline at the range of 20 deg ≤ α ≤ 30 deg. Lift coefficient variation of J2 and J3 models showed better performance than that of J1 model at the post-stall region. J1 have delayed the stall angle of attack from 10° to 11° and increased the lift coefficient by 6%. J2 and J3 show nearly the same aerodynamic characteristics, delayed the stall angle by 3 degrees compared to the baseline model and increased the lift coefficient by approximately 15%. At AOA = 15°, There is a 43% increase in lift coefficient of J2 and J3 when it compared to the baseline model. In the post-stall region, J3, having suggested waveform, indicates a slightly higher lift coefficient than J2.

Figure 6 shows the comparison of the drag coefficients for the baseline and J1, J2 and J3 modified models at Re= $1.0x10^5$. J1 model which is wavelengths (0.5c) has minimum drag. An exception of angles of attack between 18° - 30°, the drag coefficient of the baseline model is higher than those of J1, J2 and J3 modified models in all attack angles. After α = 18°, there is a similar trend in drag coefficient for baseline, J1 and J2 models.

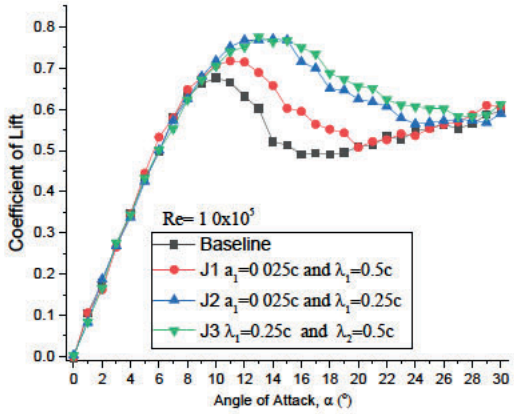


Figure 5: Comparison lift coefficients for baseline and J1, J2, J3 models at Re= $1.0x10^5$

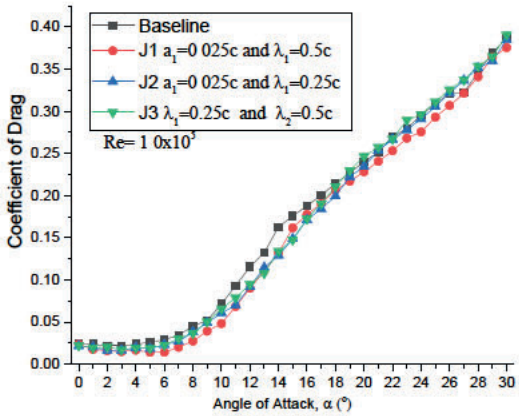


Figure 6: Comparison drag coefficients for baseline and J1, J2, J3 models at Re= $1.0x10^5$

A comparison of the lift to drag (L/D) ratio for baseline and three modified models are shown in Figure 7. It is observed that the aerodynamic performance of the modified models has a higher L/D than that of the baseline model. It is seen that the modified model, J1 provides the highest L/D ratio with its high lift and lowest drag compared to other models.

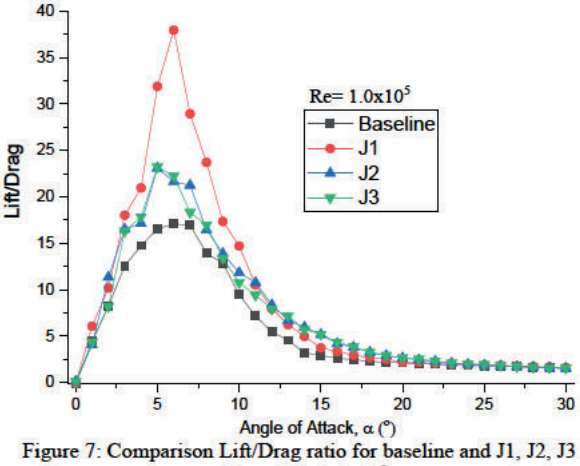


Figure 7: Comparison Lift/Drag ratio for baseline and J1, J2, J3 models at Re= $1.0x10^5$

IV. Conclusion

The effect of the tubercle leading-edge for J1, J2 and J3 models were examined experimentally at Re= $1.0x10^5$ for angles of attack between 0° - 30°. Force measurement experiments were performed for one baseline and three semi-span NACA 0020 airfoils with leading-edge tubercles. Results indicate that all the modified semi-span models improved the post-stall characteristics. Maximum drag reduction is obtained as 20.5% for the J1 model. Among the modified models, the leading-edge tubercle geometry with the best aerodynamic characteristics was determined as J1 model. In terms of lift coefficient characteristics, J3, that is suggested tubercle model, presented best lift characteristics than the other tubercle models and baseline model.

Acknowledgement

The authors would like to acknowledge the financial support of this work by the Scientific and Technological Research Council of Turkey (TUBITAK) under the Contract Number of 118M592.

References

[1] F E Fish and J M Battle, "Hydrodynamic design of the humpback whale flipper," *J. Morphol.*, vol 225, no 1, pp 51–60, 1995

[2] D S Miklosovic, M M Murray, L E Howle, and F E Fish, "Leading-edge tubercles delay stall on humpback whale (Megaptera novaeangliae) flippers," *Phys. Fluids*, vol 16, no 5, 2004, doi: 10 1063/1 1688341

[3] D S Miklosovic, M M Murray, and L E Howle, "Experimental evaluation of sinusoidal leading edges," *J. Aircr.*, vol 44, no 4, pp 1404–1408, 2007, doi: 10 2514/1 30303

[4] Ullstein Bild [retrieved 23 October 2020] from https://time com/4701442/humpback-whales-super-groups/

[5] H Johari, C Henoch, D Custodio, and A Levshin, "Effects of leading-edge protuberances on airfoil performance," *AIAA J.*, vol 45, no 11, pp 2634–2642, 2007, doi: 10 2514/1 28497

[6] K L Hansen, R M Kelso, and B B Dally, "Performance Variations of Leading-Edge Tubercles for Distinct Airfoil Profiles," *AIAA J.*, vol 49, no 1, pp 185–194, 2011, doi: 10 2514/1 J050631

[7] M L Post, A R Sapell, and J S Hart, "Bio-inspired sinusoidal leading-edged wings," *46th AIAA Fluid Dyn. Conf.*, no June, 2016, doi: 10 2514/6 2016-3255

Selection of Process Parameters on Wire Arc Additive Manufacturing(WAAM) of Ti6Al4V

E. ÇAKIR[1], İ. BAYAR[2], and M. ULUTAN[3]

[1] Eskisehir Osmangazi University, Graduate School of Natural and Applied Sciences, Eskisehir/Turkey, ersincak@gmail.com
[2] Eskisehir Osmangazi University, Department of Mechanical Engineering, Eskisehir/Turkey, ismailbayar@ogu.edu.tr
[3]Eskisehir Osmangazi University, Department of Mechanical Engineering, Eskisehir/Turkey, mulutan@ogu.edu.tr

***Abstract* – In recent years, Wire Arc Additive Manufacturing (WAAM) has remarkable development due to low equipment cost and high deposition rate. In this study, Ti6Al4V material which is widely used in the aviation industry, was deposited by Wire Arc Additive Manufacturing. Plasma arc welding machine has been used because of its high heat input which is more suitable method for titanium alloys. Investigations were carry out by changing wire feed speed, travel speed and ampere values. Graphs and tables were created by measuring layer width and layer height from single layer samples. Using these values, suitable parameters for Ti6Al4V were tried to be determined.**

***Keywords* – Additive Manufacturing, Wire Arc Additive Manufacturing (WAAM), Titanium.**

I. Introduction

Additive manufacturing, seen as a promising method of production, has become a subject of intensive study. Wire arc additive manufacturing (WAAM) has improved significantly with its low equipment cost and high deposition rates [1]. This manufacturing method; It is based on combining the wire feeding and arc welding so that parts produced layer by layer. This system consists of arc formation from a welding machine, torch, robot or manipulator to control movement, wire feeder and a base plate where the final material will be deposited [2]. Classification of the equipment used is shown in Figure 1.

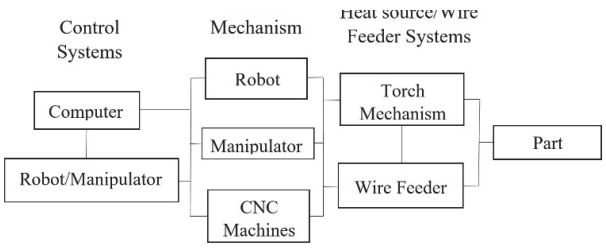


Figure 1. Classification of equipment using in wire arc additive manufacturing (WAAM)

When wire arc additive manufacturing is classified according to the types of welding used it is called by different names. Gas metal arc additive manufacturing (GMAAM) when MIG-MAG welding machine is used, gas tungsten arc additive manufacturing (GTAAM) when TIG welding machine is used and plasma arc additive manufacturing (PAAM) when Plasma arc welding machine is used. This classification is shown in Figure 2.

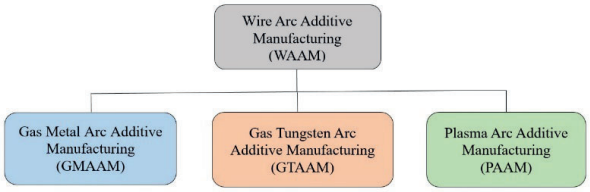


Figure 2. Wire arc additive manufacturing(WAAM) classification according to types of welding

Many of experimental studies for different material have been applied by using these methods such as steel [3], aluminum [4], nickel superalloys [5] and titanium alloys [6]. Titanium alloys, especially Ti6Al4V attract significant research interest from the proponents of additive manufacturing due to its frequent use in aerospace applications, where conventional substractive fabrication methods for these components often suffer from very low fly to buy ratios [7].

It is seen that using the Plasma arc additive manufacturing (PAAM) method, which has a high heat input, in aerospace materials that are mostly composed of titanium and its alloys, will be more efficient than other methods such as GMAAM and GTAAM [8].

This production method has become a rapidly developing and intensively studied subject with its ability to produce close to the final shape, shortening of production time, close to zero material waste and cost advantages. In addition, it is possible to manufacture complex parts in a single operation and with fewer process operations, significant reductions in production time are observed. [9,10]

In this study, parameters were determined with reference to previous studies and deposition was made for nine different parameters. For Ti6Al4V materials were tried to determine optimum parameters by using Taguchi Method.

II. EXPERIMENTAL

A. *Conditions of experiments*

In this research, Plasma Arc Additive Manufacturing (PAAM) method used and this system is contain plasma welding machine, three axis manipulator mechanism and wire feeder. This system has been developed for the experiments in Eskişehir Osmangazi University Mechanical Engineering Department.

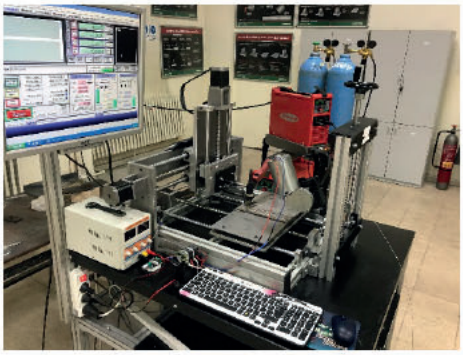


Figure 3. Wire arc additive manufacturing (WAAM) system

Ti6Al4V, 80x200x4 mm thickness substrate is selected and 1 mm diameter Ti6Al4V welding wire is used. Chemical composition of materials is presented in Table 1.
All of experiments were used 0.8 lt/min plasma flow rate and 15lt/min shielding gas flow rate. In order to determine the optimum parameters of proses, nine different parameters have been studied. Three parameters consist of ampere value, wire feed speed and travel speed were examined. These parameters are presented Table 3. The welding torch is held in a vertical position and wire is feeding front side at a fixed angle. After the deposition, all samples have been waited for cooling at a room temperature.

B. *Taguchi Method*

Taguchi is a mathematical model which can reduce the number of experiments by conducting limited effective experiments to full fill the requirement [11]. Taguchi's L9 orthogonal design which is suitable three parameters is used for experiments so Taguchi method was used for array the parameters. This L9 Taguchi's orthogonal array is shown Table 2.

Table 2. L9(3^3) Orthogonal Array of Taguchi Method [12]

Condition/Column	L_9		
	1	2	3
1	1	1	1
2	1	2	2
3	1	3	3
4	2	1	2
5	2	2	3
6	2	3	1
7	3	1	3
8	3	2	1
9	3	3	2

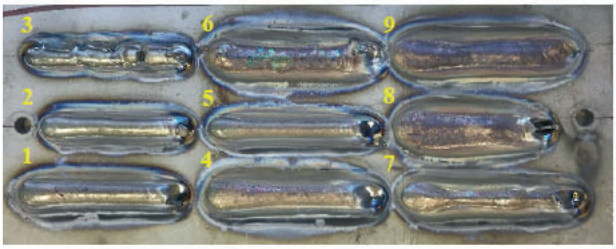


Figure 4. (a)Top views of single layer depositions

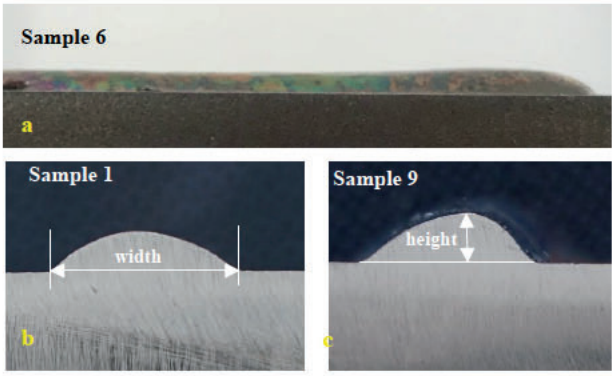


Figure 5 (a)Front view of at 90 amper, 30 mm/s wire feed speed and 2,5 mm/s travel speed, (b) Cross section of deposition at 75 amper,15 mm/s wire feed speed and 2,5 mm/s travel speed, c) Cross section of deposition at 105 amper, 30mm/s wire feed speed and 3 mm/s travel speed.

Table 1. Chemical compositions of materials in wt%.

Materials	Fe	C	N	H	O	Al	V	Ti
Ti6Al4V(Substrate)	0,09	0,012	0,018	0,0008	0,12	5,9	3,9	Balance
Ti6Al4V(Wire)	0,128	0,014	0,01	0,004	0,10	5,85	4,0	Balance

III. RESULTS AND DISCUSSION

Nine single layer samples have been deposited in this study. Approximate values were determined by taking five measurements from samples. Wall width and layer heights of samples are present in Fig 6. Layer heights are obtained 1,08.-1,95 mm and wall widths are 4,33-9,06 mm.

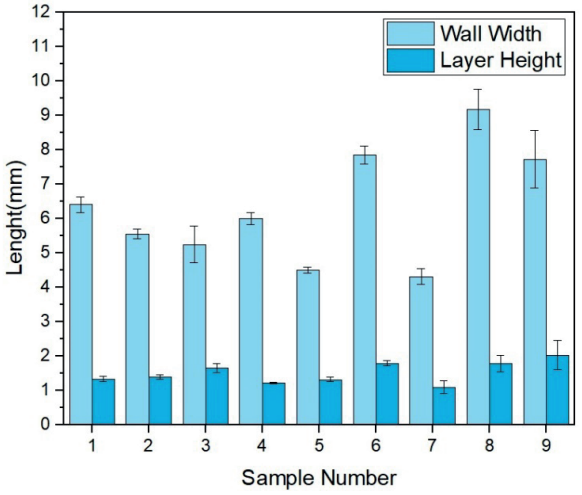


Figure 6. Layer heights and wall widths of samples

The maximum layer height were observed in the sample 9, maximum wall width were observed in sample 8 and minimum layer height and wall width observed in the sample 7.

Table 3. L9(3^3) Orthogonal array of experiments

Sample No	Curre	Wire Feeder (mm/s)	Travel Speed (mm/s)	Width(mm)	Height(mm)
1	75	15	2,5	6,36	1,32
2	75	20	3	5,49	1,39
3	75	30	4	5,24	1,65
4	90	15	3	5,97	1,23
5	90	20	4	4,51	1,33
6	90	30	2,5	7,79	1,77
7	105	15	4	4,33	1,08
8	105	20	2,5	9,06	1,75
9	105	30	3	7,85	1,95

The width and height values shown in Table 3 have been written as the average of five measurements.

The maximum height and minimum height are seen in samples 7,8 and 9 so it is seen that these values are directly related to the current. In addition, as the wire feed speed was low in sample 7, the accumulation rates were low, and in samples 8 and 9, maximum values were obtained with the increase in wire feed speed and 105 Ampere current value.

The values were measured with an electronic caliper on the samples shown in Fig. 4, and the width and height values were rounded in decimals. 5 measurements were taken from each sample. Before all measurements, the thickness of the substrate was measured precisely and the average measurement was taken as 4.05 mm since the measurement will be used at heights values.

In this experiment, in order to evaluate our variables current, wire feed speed and travel speed according to each other, graphs were created by Contour plot method according to the width of the produced material. This graphics are presented in Fig 7.

As seen in Fig.7 (a), while the width of the produced material increases in direct proportion with the increase of the current value, a significant decrease is observed in the width of the produced material as the travel speed increases over approximately 3.6 mm / s. In addition to while the wall width value of the produced material increases with the increase of the travel speed up to approximately 3mm / s, decreases in the material width are observed with the increase of the travel speed above this value at low current values.

When the effect of current and wire feed speed values on the width of the produced material is examined, it is observed that while the wire feed speed is approximately 15-18 values, the material width increases up to 95 amps, after this point, it is observed that the wall widths of the current and wire feed speed can take different values.

In examinations where current is high and wire feed speed is very low, current is low and wire feed speed is very high, it is evaluated that suitable material thickness cannot be obtained.

In general, when there is a significant decrease in the width of the produced material, increases in height are observed.

It is important that the travel speed is at a certain rate with the wire feed speed, if high wire feed speed is applied at low travel speed, very large materials can be obtained. Likewise, if low wire feed speed is applied at high travel speed, it will be observed that the width of the produced material is very low. However, these evaluations are also important to the current value.

In addition, when examining the graphics drawn by considering the wall widths, it should be taken into account that the height will change with the change of variable values.

For the production of Ti6Al4V material with the wire arc additive manufacturing; It is thought that it would be appropriate to select the parameters by evaluating the interrelationships of current, wire feeding speed and travel speed. The graphic in which this evaluation can be made easily is presented in fig 7.

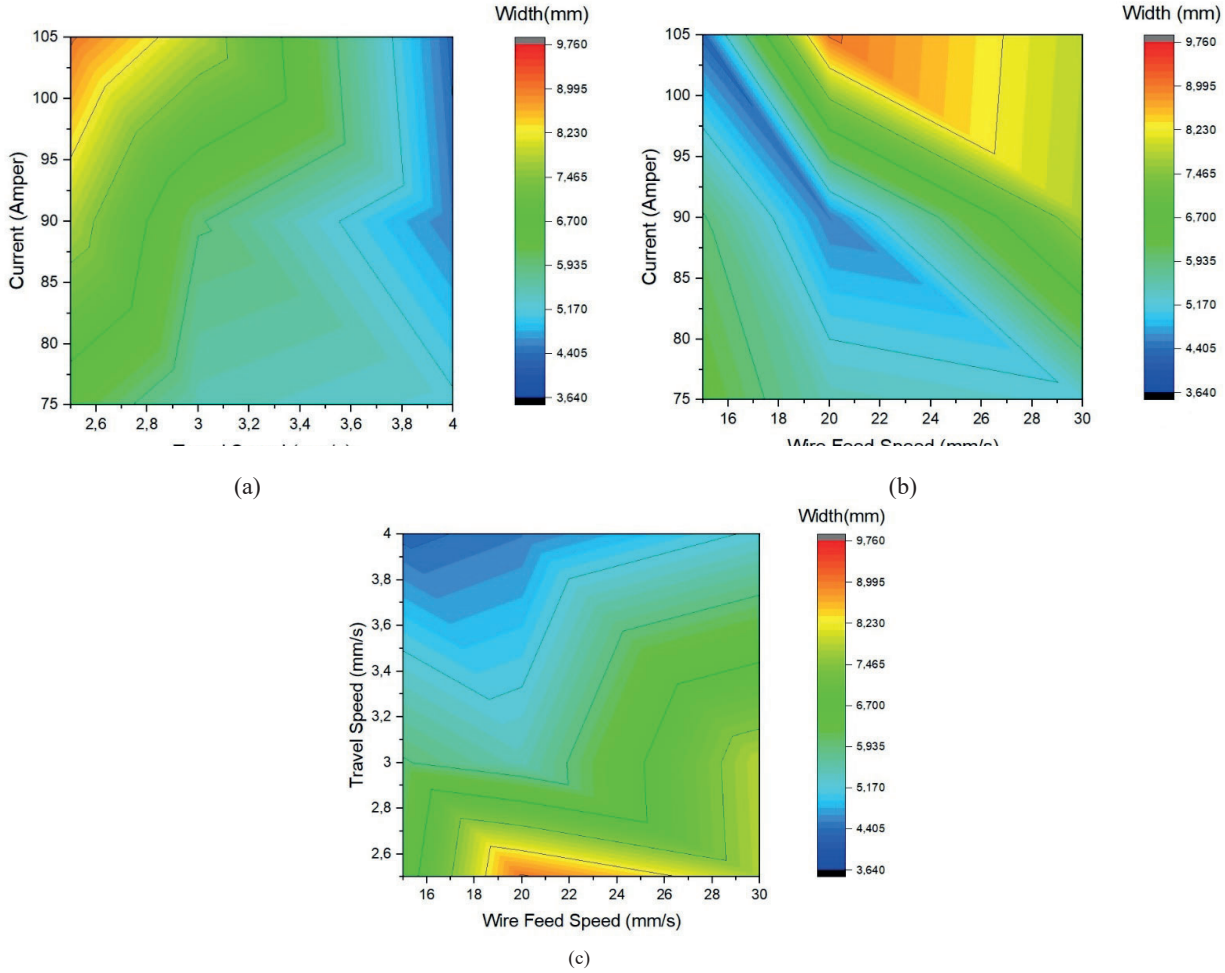


Figure 7. Contour plot for with respect to width a) TS(mm/s) and Current (ampere), b)WFS(mm/s) and Current(ampere), c) WFS(mm/s) and TS(mm/s).

IV. Conclusion

Nine samples of Ti6Al4V alloys were successfully facricated in different parameters by wire arc additive manufacturing (WAAM) in the study. In the parameters, current, wire feed speed and travel speed were selected as variables and the effects of these values on the layer height and wall width were evaluated. The following results were obtained.

- At high current values, high wall width values can be obtained with appropriate wire feeding speed.
- Increasing the travel speed causes decreases in the wall width.
- It was evaluated that it would be appropriate to increase the applied current value by increasing the wire feeding speed.

Selection of proses parameters on Wire arc additive manufacturing of Ti6Al4V alloys, it was observed that there are more suitable parameters when examining the relationship between current, wire feed speed and travel speed.

Acknowledgment

This study was supported by the Scientific Research Council of Eskişehir Osmangazi University (ESOGU-BAP), Project No: 2019-2981.

References

[1] Ivántabernero, A Paskual, P Álvarez, and A Suárez, "Study on Arc Welding Processes for High Deposition Rate Additive Manufacturing," Procedia CIRP, vol 68, no April, pp 358–362, 2018

[2] Busachi, A , Erkoyuncu, J , Colegrove, P , Martina, F , Watts, C , Drake, R , A review of Additive Manufacturing technology and cost estimation techniques for the defence sector CIRP Journal of Manufacturing Science and Technology, V:19, p:117-128, 2017

[3] M Liberini *et al.*, "Selection of Optimal Process Parameters for Wire Arc Additive Manufacturing," *Procedia CIRP*, vol 62, pp 470–474, 2017

[4] J Gu, S Yang, M Gao, J Bai, and K Liu, "Influence of deposition strategy of structural interface on microstructures and mechanical properties of additively manufactured Al alloy," Addit Manuf , vol 34, no January, p 101370, 2020
[5] M Benakis, D Costanzo, and A Patran, "Current mode effects on weld bead geometry and heat affected zone in pulsed wire arc additive manufacturing of Ti-6-4 and Inconel 718," *J. Manuf. Process.*, vol 60, no July, pp 61–74, 2020
[6] K Wang, Y Liu, Z Sun, J Lin, Y Lv, and B Xu, "Microstructural evolution and mechanical properties of Inconel 718 superalloy thin wall fabricated by pulsed plasma arc additive manufacturing," *J. Alloys Compd.*, vol 819, pp 30–40, 2020
[7] B Wu, Z Pan, D Ding, D Cuiuri, H Li, and Z Fei, "The e ff ects of forced interpass cooling on the material properties of wire arc additively manufactured Ti6Al4V alloy," *J. Mater. Process. Tech.*, vol 258, no March, pp 97–105, 2018
[8] J Lin *et al.*, "Journal of the Mechanical Behavior of Biomedical Materials Microstructural evolution and mechanical property of Ti-6Al-4V wall deposited by continuous plasma arc additive manufacturing without post heat treatment," vol 69, no November 2016, pp 19–29, 2017
[9] A A Antonysamy, J Meyer, P B Prangnell, Effect of build geometry on the β-grain structure and texture in additive manufacture of Ti-6Al-4V by selective electron beam melting, Mater Charact 84, 153–168, 2013
[10] W E Frazier, Metal additive manufacturing: a review, J Mater Eng Perform 23 1917–1928, 2014
[11] A Manjunath, V Anandakrishnan, S Ramachandra, and K Parthiban, "Experimental investigations on the effect of pre-positioned wire electron beam additive manufacturing process parameters on the layer geometry of titanium 6Al4V," Mater Today Proc , vol 21, pp 766–772, 2020
[12] R Roy, A Primer on the Taguchi Method: Society of Manufacturing Engineer, 2010

Energy and Exergy Analyses of a Combined Power System Utilizing Cold Energy of LNG

D. ÇALIKIRAN[1], S. KIRÇİÇEK[1], D. N. OZEN[1] and M. E. TOLU[2]

[1]Necmettin Erbakan University, Konya/Turkey, dnozen@konya.edu.tr
[2]Karamanoglu Mehmetbey University, Karaman/Turkey, metolu@kmu.edu.tr

Abstract **- In this study, a combined power system is aimed to utilize from the cold energy of LNG. The combined power system consists of a 2-stage ORC cycle and a direct expansion LNG cycle. A mathematical model was created for the analysis of the system. Key parameters for energy and exergy analysis were determined as propane mass ($\dot{m}_{pro}$), pressure ratio (P_r) and heat exchanger-3 outlet temperature (T_8), and the effect of these parameters on system performance was investigated for certain intervals. It was observed that as $\dot{m}_{pro}$, P_r and T_8 increased, the thermal and exergy efficiency of the combined power system increased.**

Keywords **– LNG cold energy, ORC cycle, energy, exergy, combined power system.**

I. Introduction

TRANSPORTING natural gas, which is one of the important energy resources, over long distances is preferred in LNG form due to economic and geographic reasons. Natural gas must be cooled down to -162 ° C while it is extracted to make it portable. LNG must be re-gasified in order to be used as fuel. Cooling energy is generated in this process. Rearranging LNG requires about 800 kJ / kg of heat energy [1].

There are many studies in the literature on power cycles using the cold energy of liquefied natural gas. Badami and Bruno [2] analyzed three combined cycle systems for power generation: the first is a direct expansion cycle with a Rankine cycle, the second offers double expansion with reheat and a recovery heat exchanger, and the last one is under two different turbine pressures running parallel Rankine cycles. The performance of the three cycles was compared and the effects of using work fluids with different characteristics were analyzed in detail. The main goal of Sadreddini and Ashjari [3] is to find the best cycle to recover energy from two different sources, including a medium temperature wastewater, heat and water vapor cooling. For this, they considered three different cycles and compared the optimum performances. These systems include a transcritical carbon dioxide cycle, an ORC, and a combination of the two. Bao and Yuan [4] compared eight different systems by combining one-stage Rankine cycle and two-stage Rankine cycle in series and parallel, respectively. Taking the net power output as an objective function, they optimized the basic parameters and working fluids of the eight systems at four different LNG evaporation pressures simultaneously.

In this study, energy and exergy analysis of a new combined power system using sea water as a heat source and cold energy of LNG as a heat well has been performed.

II. System Description

The proposed combined power system shown in Figure 1 consists of an ORC cycle and a direct expansion LNG cycle.

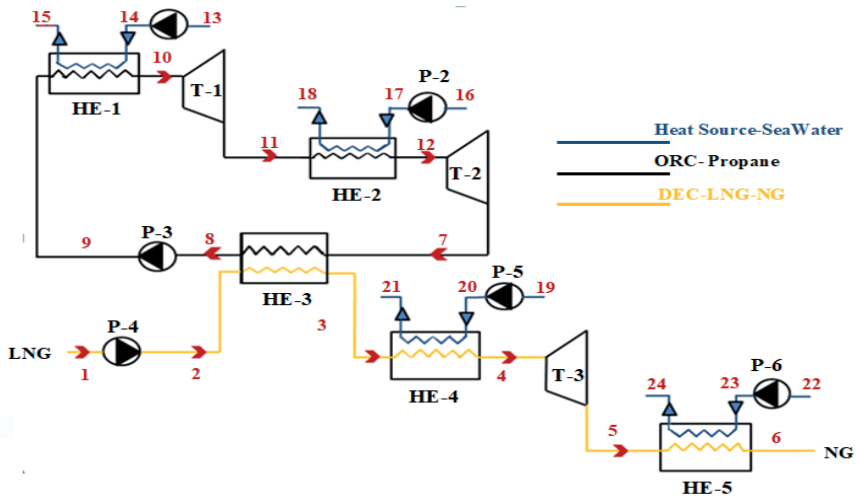


Figure 1: Schematic diagram of proposed combined power system

In the ORC cycle, sea water is used as the heat source and the cold energy of LNG is used as the place where heat is discharged. In addition, a direct expansion LNG cycle is included in a combined power system in order to benefit from this energy of pressurized natural gas converted into gas phase.

III. Mathematical Modelling of Proposed Combined Power System

A mathematical model is created for the energy and exergy analysis of the combined power system. The following assumptions have been made for the analysis of system performance.

- Combined power system operates in steady state condition.
- Kinetic and potential energy changes are neglected.
- The turbine, pump in the combined system has isentropic efficiency.
- Ambient temperature and pressure are assumed to be 298 K and 100 kPa, respectively.

The first input acceptance of some parameters for the solution of the mathematical model and the design of the system are presented in Table 1.

Table 1: Design input parameters

Parameter	Symbol	Value
Sea Water Line		
Sea water pump inlet pressure	P_{13} P_{16} P_{19} P_{22}	100 (kPa)
Sea water pump inlet temperature	T_{13} T_{16} T_{19} T_{22}	20(ºC)
Sea water pump outlet temperature	T_{15} T_{18} T_{21} T_{24}	20(ºC)
Organic Rankine Cycle		
Pressure rate	P_r	25
Propane flow rate	$\dot{m}_{pro}$	110(kg/s)
Superheating temperature	T_{sh}	3(ºC)
Condenser temperature	T_8	-70 (ºC)
Turbine isentropic efficiency	$\eta_{T-1,2}$	0.90
Pump isentropic efficiency	η_{P-3}	0.85
LNG Line		
LNG flow rate	$\dot{m}_{LNG}$	100 (kg/s)
Turbine 3 inlet pressure	P_4	8000(kPa)
Natural gas distribution temperature	T_6	0 (ºC)

A. Energy Analysis

Energy balance equations for each component of the combined power system are given in Table 2.

Table 2: Energy balance equations for the components of the system

Component	Energy Balance Equation	
ORC		
T-1	$\dot{W}_{T-1} = \dot{m}_{pro}(h_{10} - h_{11})$	$\eta_{T-1} = \frac{(h_{10} - h_{11})}{(h_{10} - h_{11s})}$
T-2	$\dot{W}_{T-2} = \dot{m}_{pro}(h_{12} - h_7)$	$\eta_{T-2} = \frac{(h_{12} - h_7)}{(h_{12} - h_{7s})}$
HE-1	$\dot{Q}_{HE-1} = \dot{m} \quad (h_{10} - h_9)$	
HE-2	$\dot{Q}_{HE-2} = \dot{m} \quad (h_{12} - h_{11})$	
HE-3	$\dot{Q}_{HE-3} = \dot{m} \quad (h_7 - h_8)$	
P-1	$\dot{W}_{P_{-1}} = \dot{m}_{sw1}(h_{14} - h_{13})$	$\eta_{P-1} = \frac{(h_{14s} - h_{13})}{(h_{14} - h_{13})}$
P-2	$\dot{W}_{P_{-2}} = \dot{m}_{sw2}(h_{17} - h_{16})$	$\eta_{P-2} = \frac{(h_{17s} - h_{16})}{(h_{17} - h_{16})}$
P-3	$\dot{W}_{P-3} = \dot{m}_{pro}(h_9 - h_8)$	$\eta_{P-3} = \frac{(h_{9s} - h_8)}{(h_9 - h_8)}$
LNG Line		
T-3	$\dot{W}_{T-3} = \dot{m}_{LNG}(h_4 - h_5)$	$\eta_{T-3} = \frac{(h_4 - h_5)}{(h_4 - h_{5s})}$
HE-4	$\dot{Q}_{HE-4} = \dot{m}_{LNG}(h_4 - h_3)$	
HE-5	$\dot{Q}_{HE-5} = \dot{m}_{LNG}(h_6 - h_5)$	
P-4	$\dot{W}_{P-4} = \dot{m}_{LNG}(h_2 - h_1)$	$\eta_{P-4} = \frac{(h_{2s} - h_1)}{(h_2 - h_1)}$
P-5	$\dot{W}_{P_{-5}} = \dot{m}_{sw3}(h_{20} - h_{19})$	$\eta_{P4} = \frac{(h_{20s} - h_{19})}{(h_{20} - h_{19})}$
P-6	$\dot{W}_{P_{-6}} = \dot{m}_{sw4}(h_{23} - h_{22})$	$\eta_{P4} = \frac{(h_{23s} - h_{22})}{(h_{23} - h_{22})}$

The net power and thermal efficiency of the system are found with the following equations.

$$\dot{W}_{ORC} = \dot{W}_{T1} + \dot{W}_{T2} - (\dot{W}_{P1} + \dot{W}_{P2} + \dot{W}_{P3}) \quad (1)$$

$$\dot{W}_{LNG} = \dot{W}_{T3} - (\dot{W}_{P4} + \dot{W}_{P5} + \dot{W}_{P6}) \quad (2)$$

$$\dot{W}_{net} = \dot{W}_{ORC} + \dot{W}_{LNG} \quad (3)$$

$$\dot{Q}_{in} = \dot{Q}_{HE1} + \dot{Q}_{HE2} + \dot{Q}_{HE4} \quad (4)$$

Thermal efficiency of the system:

$$\eta = \frac{\dot{W}_{NET}}{\dot{Q}_{in}} \quad (5)$$

B. Exergy Analysis

Fuel and product exergy equations and exergy destruction equations for each component of the combined system are given in Table 3. Kinetic and potential exergy have been neglected.

Table 3: Fuel and product exergy definitions for the system components

Component	Fuel	Product
P-1	$\dot{W}_{P-1}$	$\left(\dot{Ex}^{T}_{14}-\dot{Ex}^{T}_{13}\right)+\left(\dot{Ex}^{M}_{14}-\dot{Ex}^{M}_{13}\right)$
HE-1	$\left(\dot{Ex}^{T}_{14}+\dot{Ex}^{T}_{9}\right)+\left(\dot{Ex}^{M}_{14}+\dot{Ex}^{M}_{9}\right)-\left(\dot{Ex}^{M}_{10}+\dot{Ex}^{M}_{15}\right)$	$\left(\dot{Ex}^{T}_{10}+\dot{Ex}^{T}_{15}\right)$
T-1	$\left(\dot{Ex}^{T}_{10}-\dot{Ex}^{T}_{11}\right)+\left(\dot{Ex}^{M}_{10}-\dot{Ex}^{M}_{11}\right)$	$\dot{W}_{T-1}$
P-2	$\dot{W}_{P-2}$	$\left(\dot{Ex}^{T}_{17}-\dot{Ex}^{T}_{16}\right)+\left(\dot{Ex}^{M}_{17}-\dot{Ex}^{M}_{16}\right)$
HE-2	$\left(\dot{Ex}^{T}_{17}+\dot{Ex}^{T}_{11}\right)+\left(\dot{Ex}^{M}_{17}+\dot{Ex}^{M}_{11}\right)-\left(\dot{Ex}^{M}_{18}+\dot{Ex}^{M}_{12}\right)$	$\left(\dot{Ex}^{T}_{18}+\dot{Ex}^{T}_{12}\right)$
T-2	$\left(\dot{Ex}^{T}_{12}-\dot{Ex}^{T}_{7}\right)+\left(\dot{Ex}^{M}_{12}-\dot{Ex}^{M}_{7}\right)$	$\dot{W}_{T-2}$
HE-3	$\left(\dot{Ex}^{T}_{7}+\dot{Ex}^{T}_{2}\right)+\left(\dot{Ex}^{M}_{7}+\dot{Ex}^{M}_{2}\right)-\left(\dot{Ex}^{M}_{8}+\dot{Ex}^{M}_{3}\right)$	$\left(\dot{Ex}^{T}_{8}+\dot{Ex}^{T}_{3}\right)$
P-3	$\dot{W}_{P-3}$	$\left(\dot{Ex}^{T}_{9}-\dot{Ex}^{T}_{8}\right)+\left(\dot{Ex}^{M}_{9}-\dot{Ex}^{M}_{8}\right)$
P-4	$\dot{W}_{P-4}$	$\left(\dot{Ex}^{T}_{2}-\dot{Ex}^{T}_{1}\right)+\left(\dot{Ex}^{M}_{2}-\dot{Ex}^{M}_{1}\right)$
P-5	$\dot{W}_{P-5}$	$\left(\dot{Ex}^{T}_{20}-\dot{Ex}^{T}_{19}\right)+\left(\dot{Ex}^{M}_{20}-\dot{Ex}^{M}_{19}\right)$
HE-4	$\left(\dot{Ex}^{T}_{20}+\dot{Ex}^{T}_{3}\right)+\left(\dot{Ex}^{M}_{20}+\dot{Ex}^{M}_{3}\right)-\left(\dot{Ex}^{M}_{21}+\dot{Ex}^{M}_{4}\right)$	$\left(\dot{Ex}^{T}_{21}+\dot{Ex}^{T}_{4}\right)$
T-3	$\left(\dot{Ex}^{T}_{4}-\dot{Ex}^{T}_{5}\right)+\left(\dot{Ex}^{M}_{4}-\dot{Ex}^{M}_{5}\right)$	$\dot{W}_{T-3}$
P-6	$\dot{W}_{P-6}$	$\left(\dot{Ex}^{T}_{23}-\dot{Ex}^{T}_{22}\right)+\left(\dot{Ex}^{M}_{23}-\dot{Ex}^{M}_{22}\right)$
HE-5	$\left(\dot{Ex}^{T}_{23}+\dot{Ex}^{T}_{5}\right)+\left(\dot{Ex}^{M}_{23}+\dot{Ex}^{M}_{5}\right)-\left(\dot{Ex}^{M}_{24}+\dot{Ex}^{M}_{6}\right)$	$\left(\dot{Ex}^{T}_{24}+\dot{Ex}^{T}_{6}\right)$

Chemical exergy has been taken into account only in the combustion chamber, since no chemical reaction occurs in the other components. Physical and chemical exergy were calculated from the following equations.

$$e_{ph} = h - h_o - T_o(s - s_o) \quad (6)$$

$$\dot{E}_{in} = \dot{E}_{F,HE1} + \dot{E}_{F,HE2} + \dot{E}_{F,HE4} + \dot{E}_{F,HE5} \quad (7)$$

Exergy efficiency of the system:

$$\Psi = \frac{\dot{W}_{NET}}{\dot{E}_{in}} \quad (8)$$

IV. Results and Discussion

It has been studied how the change of key parameters affects system performance. The key parameters are the propane mass ratio, the pressure ratio in the ORC cycle and the outlet temperature of the heat exchanger 3. The effects of these parameters on the energy and exergy efficiencies and the exergy efficiency were investigated (Figure 2). System performance is compared by changing the design input values. In the parametric study, it should be noted that when one variable changes, the values in Table 1 of the other variables are kept constant. Thermodynamic properties are calculated for each component of the system according to the design input parameters given in Table 1 and are presented in Table 4.

Table 4: Thermodynamic data for the proposed system

State	$\dot{m}$ (kg/s)	T (℃)	P (kPa)	$\dot{E}x^T$(kW)	$\dot{E}x^M$ (kW)	$\dot{E}x$ (kW)
1	100	-162	100	108404	0	108404
2	100	-159	8000	44086	65644	109730
3	100	-48.41	8000	4309	65644	69953
4	100	0	8000	326.50	65644	65970.5
5	100	-44.07	4000	2527	55943	58470
6	100	0	4000	276.30	55943	56219.30
7	110	-36.38	24.34	1205	-8660	-7455
8	110	-70	24.34	26286	-8660	17626
9	110	-69.76	608.50	7084	10639	17723
10	110	11.41	608.50	63.73	10639	10702.73
11	110	-37.86	121.70	1468	1192	2660
12	110	11.41	121.70	57.04	1192	1249.04
13	1443	20	100	240.60	0	240.60
14	1443	20.01	300	239.40	289.40	528.80
15	1443	10	300	2310	289.40	2599.40
16	215.80	20	100	35.98	0	35.98
17	215.80	20.01	300	35.81	43.28	79.09
18	215.80	10	300	345.60	43.28	388.88
19	446.90	20	100	74.52	0	74.52
20	446.90	20.01	300	74.15	89.63	163.78
21	446.90	10.0000	300	715.60	89.63	805.23
22	280	20	100	46.68	0	46.68
23	280	20.01	300	46.46	56.15	102.61
24	280	10	300	448.30	56.15	504.45
$\dot{W}_{P-1}$	-	-	-	-	-	340.1
$\dot{W}_{T-1}$	-	-	-	-	-	7050
$\dot{W}_{P-2}$	-	-	-	-	-	50.86
$\dot{W}_{T-2}$	-	-	-	-	-	7625
$\dot{W}_{P-3}$	-	-	-	-	-	123.4
$\dot{W}_{P-4}$	-	-	-	-	-	2186
$\dot{W}_{P-5}$	-	-	-	-	-	105.3
$\dot{W}_{T-3}$	-	-	-	-	-	6549
$\dot{W}_{P-6}$	-	-	-	-	-	65.99

As seen in Figure 2, the energy and exergy efficiencies increase with the increase of propane mass ratio, pressure ratio in the ORC cycle and the exit temperature of the heat exchanger 3.

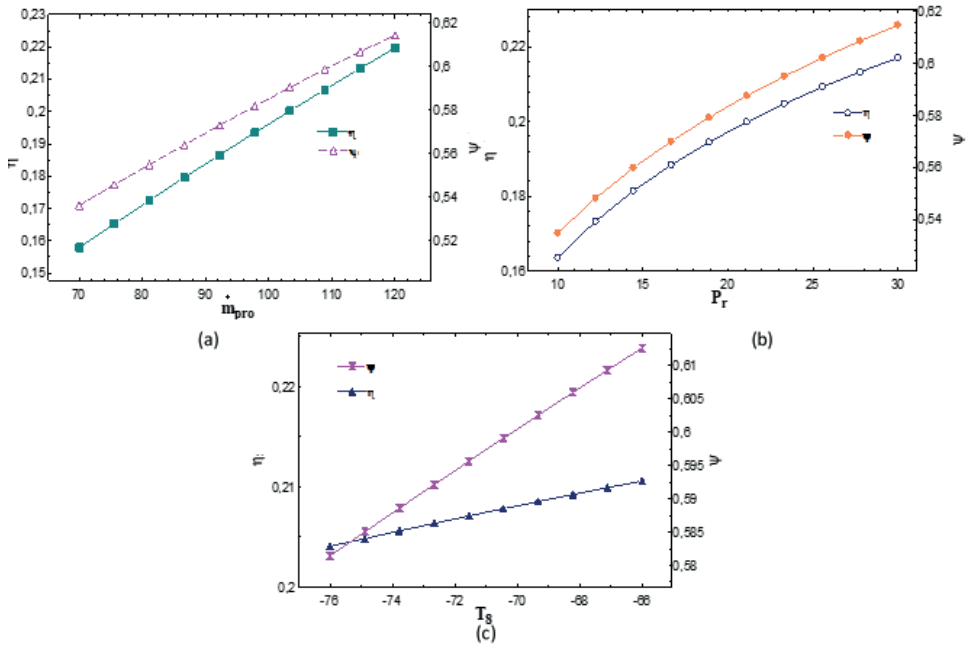


Figure 2: The effects of propane mass ratio (a), the pressure ratio in the ORC cycle (b) and the outlet temperature of the heat exchanger 3 (c) on energy and exergy efficiencies

V. Conclusion

In this study, a new combined power system using LNG's cold energy is designed. The effect of selected key parameters on the performance of the system was investigated. It was revealed that the increase in the propane mass ratio, the pressure ratio in the ORC cycle and the exit temperature of the heat exchanger 3 increased the energy and exergy efficiency.

VI. References

[1] A Franco, and C Casarosa, "Thermodynamic analysis of direct expansion configurations for electricity production by LNG cold energy recovery," *Applied Thermal Engineering*, vol 78, pp 649-657, 2015

[2] M Badami, J C Bruno, A Coronas, and G Fambri, "Analysis of different combined cycles and working fluids for LNG exergy recovery during regasification," *Energy*, vol 159, pp 373-384, 2018

[3] A Sadreddini, M A Ashjari, M Fani, and A Mohammadi, "Thermodynamic analysis of a new cascade ORC and transcritical CO2 cycle to recover energy from medium temperature heat source and liquefied natural gas," *Energy Conversion and Management*, vol 167, pp 9-20, 2018

[4] J Bao, T Yuan, L Zhang, N Zhang, X Zhang, and G He, "Comparative study of liquefied natural gas (LNG) cold energy power generation systems in series and parallel," *Energy conversion and management*, vol 184, pp 107-126, 2019

[5] S Lee, "Multi-parameter optimization of cold energy recovery in cascade Rankine cycle for LNG regasification using genetic algorithm," *Energy*, vol 118, pp 776-782, 2017

[6] X Peng, X She, B Nie, C Li, Y Li, and Y Ding, "Liquid Air Energy Storage with LNG cold recovery for air liquefaction improvement," *Energy Procedia*, vol 158, pp 4759-4764, 2019

[7] M R Gómez, R F Garcia, J R Gómez, and J C Carril, "Thermodynamic analysis of a Brayton cycle and Rankine cycle arranged in series exploiting the cold exergy of LNG (liquefied natural gas)," *Energy*, vol 66, pp 927-937, 2014

[8] M R Gómez, R F Garcia, J R Gómez, and J C Carril, "Review of thermal cycles exploiting the exergy of liquefied natural gas in the regasification process," *Renewable and Sustainable Energy Reviews*, vol 38, pp 781-795, 2014

Investigation of Effects of Boron Additives on Performance of Cam Mechanisms

Hasan BAŞ[1]

[1] Karadeniz Technical University, Trabzon/Turkey, bas@ktu.edu.tr

Abstract **- The use of additives in the contacting surfaces is essential for the formation of a thin lubricant film, which separates the moving surfaces from each other. Boron compounds, in particular, are very helpful for the sliding surfaces. The aim of the investigation is to clarify how the boron additives influence the lubricating ability of the engine oil using in the sliding contact of the cam-follower. In this study, the considerable results in the tests of engine oil with boron compounds and with base oil are obtained. And also, the influences of the minimize interface friction has been observed under different test conditions.**

Keywords **- Oil Additives, Internal Combustion Engine, Lubricant, Boron Compounds, Friction Force, Friction Coefficient, Boric Acid, Boron Nitride.**

I. Introduction

In many applications like gears, cams, valves, the operating conditions warrant the operation to lie in the mixed lubrication regime[1]. Besides, under very heavy loads, slow relative speed, insufficient surface area and scarcity of lubricant, the formation of the thick film necessary for hydrodynamic lubrication become difficult and interacting surfaces contact each other at several locations and tribo–pairs operate in mixed lubrication. The basic principle of lubrication is to keep an oil film between separated surfaces of the machine part moving relative to each other and to minimize wear. Additives in the lubricant strongly influence the behavior and the operational lifetime of the cam mechanisms. [2]. Ideally, this would be accomplished with an oil film thick enough to eliminate any contact in all internal combustion engine parts and bearings, but this may not be always possible. Therefore, to provide new desirable properties, improve existing features, and eliminate some undesirable features or at least reduce them to a minimum level, some additives are added to the lubricants [3]. In an actual engine, the lubricating oil film thickness will range from a continuous film of oil in hydrodynamic lubrication, as in crankshaft bearings, to a trapped and thinner film in elastohydrodynamic lubrication and boundary or mixed lubrication [4], as in cam-follower mechanisms. Hence, to improve the lubricating ability of the base oil and to make motor oil more slippery, boron-based additives are used with base oil which reduce friction between engine parts in an automobile and thus increasing wear resistance [5-7].

Boron compounds have a unique combination of tribological properties and unique ways of interacting with sliding surfaces to provide low friction and wear under both dry and lubricated sliding conditions. In fact, several journal articles provide information on the properties of certain solid lubricants that were mixed with a range of liquid and greases lubricants [8, 9-11]. Without the boric acid particles, the friction coefficient of base oil is around ~0.15. When base oil is blended with nanoscale boric acid particles, the friction coefficient is reduced to 0.04. Because of their layered structure, they can shear easily to provide low friction [8]. Hexagonal boron nitride is a well-known lamellar solid lubricant like graphite [8,12, 13]. In its atomic layers, boron and nitrogen are covalently bonded to one another to form a two-dimensional atomic sheet, while bonding between the layers is primarily van der Waals type and weak. The hBN has self-lubricity and can be used for mechanical seals and high-temperature solid lubricants and as padding between metal and ceramic to make bearings. Kimura and his coworkers showed that the addition of hBN powders into oil at concentrations as low as 1 wt% significantly reduced wear [7]. Hexagonal boron nitride (hBN) and boric acid stand out as very effective solid lubricants. They are very soft, but due to their unique lamellar structures, they provide some of the lowest friction coefficients in a variety of tribological applications [7,8,14]. In addition, they are generally believed to have little if any detrimental effect on either the machine components or the environment [7]. Specifically, boron compounds have been extensively studied as boundary-lubricating oil-soluble additives, solid lubricants, and surface coatings [15,16,17]. Therefore, research on boron-based additives has considerably increased within the last decade. But there are little data currently available on the tribological behavior of base engine oil and boron additives on the cam mechanisms. Thus, two types of boron compounds and their effect on the base engine oil were investigated using the cam and follower mechanism, designed and manufactured in the laboratory. The results can guide the tribological application of engine oil using with different boron compound additives on the cam mechanisms, existing in internal combustion engines and many industrial applications.

II. Theoretical Bases

A. Force Analysis

In the cam mechanism, the contact load acting on the cam-follower contact area varies depending on the cam rotation angle. Maximum contact load occurs at cam angles where the cam nose is in contact with the follower.

The loading state of the flat-surface follower-cam pair is shown in Figure 1. Here, the contact load is in the same direction as the centerline during a cam period due to the flatness of the

follower surface. The contact load at the contact surface is constant only along the base circle. The cam surface acts on the follower surface with the cam pressure angle depending on the cam rotation angle. The maximum value of the pressure angle in question occurs in the cam position where the contact point of the two surfaces is furthest from the centerline, as shown in the Fig.1. Thus, the cam remains under the influence of the maximum torque at the rotation angles corresponding to these positions.

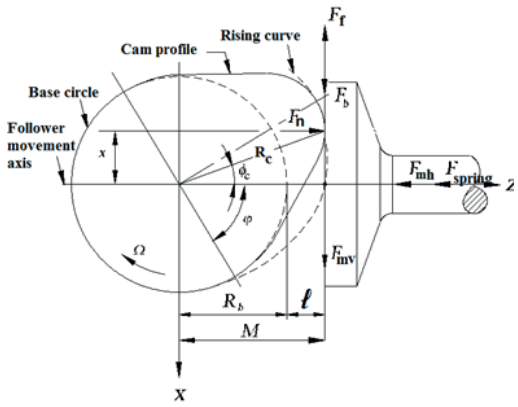


Figure 1: Loading status of flat-surfaced cam-follower pair.

B. Friction Force

The regime of true mixed lubrication, where there are both elastohydrodynamic and boundary lubricated regions within a single contact, remains quite poorly understood. The cam-follower pair is a very complex lubricated contact because of its continuous variation of load, speed and radius of curvature. In order to determine the friction force arising in the elastohydrodynamic contact under the forces acting through a period in the cam mechanism, the calculation process was made as follows.

$$F_s = \ell(u_1 - u_2)\int_{-b}^{b} \frac{\eta}{h_{cen}} dx \qquad (1)$$

Where, ℓ is cam stroke, u_1 and u_2 are velocity components at the x direction of the contact points of the cam and the follower, is lubricant viscosity, and is given in Equation 2

$$= \eta_0 e^{\alpha p} \qquad (2)$$

and hertz pressure distribution at the contact is given with Equation 3

$$p_{(x)} = p_{(\max)}\sqrt{1-\left(\frac{x}{b_{Hz}}\right)^2} \qquad (3)$$

By substituting Equations (2) and (3) in Equation (1), the friction force on the contact of the cam follower can be written as follows.

$$F_s = \frac{\eta_0(u_1-u_2)}{h_{cen}} \cdot \ell \cdot \int_{-b}^{b} e^{\alpha \cdot p_{\max}\sqrt{\left(1-\frac{x^2}{b^2}\right)}} dx \qquad (4)$$

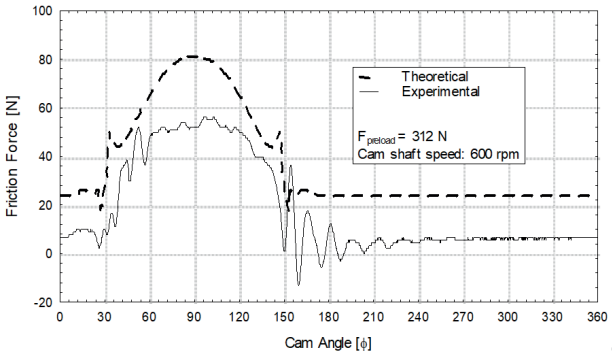


Figure 2: Experimental and theoretical results of the friction force at 312N contact load.

Where, h_{cen} is the film thickness at the lubricated contact between cam and follower where the maximum hertz pressure occurs [18]. The integration process encountered in the calculation of the friction force is solved numerically. It is seen that the results obtained in theoretical and experimental studies are in harmony with each other. However, a slight difference has been observed in the theoretical and experimental friction force graphs, for example, as seen in Figure 2. The reason for this is that the Barus equation, which gives the pressure-viscosity relationship in the equation (4) used in the theoretical analysis, gives more values than the actual values of the lubricant behavior of non-conformal contacts under heavy loading conditions [19]. Therefore, many rheological models have been put forward to determine lubricant behavior in elastohydrodynamic contact in lubricant rheology. It has been observed that there are also differences between the theoretical and experimental friction coefficient results obtained from these models [20].

III. Experimental Setup

A. Description of The Test Rig

The mechanical system and measurement unit of the test apparatus is designed to measure the frictional behavior between cam and follower at different loads, speeds and lubricating conditions. A lubricating system is designed to investigate the effect of different type of oils on the friction behavior at the cam/ follower interface. In order to ensure safe and accurate data collection and storage, a computer and a data acquisition system are attached to the measurement unit and a computer program is written to evaluate this data.

The test apparatus was manufactured in the Mechanical Engineering Department at Karadeniz Technical University. Detailed information about the test apparatus can be obtained from Ref. [21]. The test apparatus mainly consists of a driving unit, a measurement unit, a cam/ follower system, a loading unit and a lubricating unit as illustrated schematically in Ref. [21].

The camshaft is driven by a 5.5 kW induction type electric motor whose speed can be changed by a speed control unit in the range of 0–1200 rpm. There is a belt-pulley mechanism between the cam shaft and the electric motor and a flywheel is added to the driving shaft to provide a stable rotation of the cam. The flywheel is supported by two ball bearings and connected to the cam shaft via a gear type clutch. The cam shaft consists of a 32 mm-base diameter cam profile and a 7.888 mm-cam stroke which are actual case and is held rigidly to the stationary plate by two journal bearings. A measurement beam for calculating friction force is mounted on a carriage pad, which is rigidly connected to the cylinder. The follower is fixed vertically to the measurement beam and applies a preload to the cam surface by means of a spring. The loading of the follower is provided by the spring stroke, which changes via the rotation of cam. The pairs of cam and follower were selected from Mercedes cam and INA follower.

B. Measurement Procedures

To determine the friction force, the balance of forces applied on the follower in the vertical direction should be considered. The friction force applied to the follower by cam, F_f, causes a tensile effect on the measurement beam (F_b) and the effect produces a signal proportional to the beam force, which can be used in the calculation of the friction force. The balance of forces in the vertical direction applied on the follower can be written as:

$$F_f = F_{mv} + \mathrm{Fb} \tag{5}$$

where F_{mv} is the mass force of the follower that can be determined by using the following equation:

$$F_{mv} = m_v \cdot a_v \tag{6}$$

where ''*mv*'' and ''*av*'' are mass and acceleration of follower in the vertical direction, respectively. The friction coefficient can then be easily calculated from the well-known Coulomb friction law as:

$$\mu = F_f / F_n \tag{7}$$

$$F_n = F_{spring} + F_{mh} \tag{8}$$

Where F_n is the normal force applied on the contact region of the cam/ follower, F_{spring} is the spring force and F_{mh} is the mass force arising from the moving masses in the horizontal direction. The mass force can be calculated as ($F_{mh} = m_h \cdot a_h$). Where "m_h" is the mass of moving parts in the horizontal direction and "a_h" is the acceleration. The spring force is, then, calculated from a calibrated spring in terms of the cam stroke. The spring force includes the preloading force and the stroke force of the spring acquired by the reduction in spring length due to cam movement. In this system, a spring having a spring coefficient of 104.7 N/mm is used and the calibration equation is Force [N] =104.7·x displacement of spring (mm).

In the test apparatus, two types of accelerometers were used to measure the mass force in vertical and horizontal directions due to the acceleration of moving masses. The mass force in the horizontal direction is measured by using a piezoelectric type of accelerometer sensor, which works by a compression effect and is mounted rigidly between two linear bearings in the loading system. This acceleration meter has a voltage sensitivity of 7.43 V/ms^{-2} and strain sensitivity of 0.003 ms^{-2} /µstrain in the horizontal direction. An acceleration sensor is used for the calculation of friction force in the vertical direction in order to measure the mass force and mounted below the follower. This sensor has an accuracy factor of ≤0.03. The signals taken from the sensor are amplified by a charge amplifier and are sent to a computer via a data acquisition system. Temperatures of the test samples, on the other hand, were measured with K type thermocouple. The end of thermocouple connected to specimen, quick and reliable contact temperature readings via functional millivoltmeter is possible.

C. Base Oil and Boron Additives

Typical physical characteristics of the base oil (20W50 engine oil), used in this study, are given in Table 1. As additives, two kinds of boron compounds hexagonal boron nitride and boric acid powders used. Because of the additive content in the base oil was low, the viscosity variation of oil was neglected.

Table 1: Typical physical characteristics of the base oil

SAE Viscosity Grade		Base Oil (20W50)
Kinematic viscosity @ 40°C cSt 100°C cSt	ASTM D 445	688.8
Dynamic viscosity @ −25°C mPa s −15°C mPa s	ASTM D 5293	— 8.910
Viscosity Index	ASTM D 2270	102
Density @ 15°C kg/L	ASTM D 4052	0.890
Flash Point COC, °C	ASTM D 92	236
Pour Point, °C	ASTM D 97	−21

Note. COC = Cleveland Open Cup

D. Test Methodology

To investigate the effects of boron compound additives on the motor base oil, and to determine their optimum ratios, base oil was prepared by addition of 2wt.%, 4wt.%, 6wt.% HBN and BA and the optimum concentration ratio was determined as 4 wt% for boric acid (H3BO3) and 4 wt% for hexagonal boron nitride (hBN) [2]. So, in this study, experiments carried out by addition 4wt HBN and BA to the base oil. Samples were tested at various loads and speeds under lubricated conditions. Experiments were conducted at 400, 600 and 800 rpm, and 212, 312 412 N contact loads. Experimental periods were selected as 10 minutes for every test after running-in, and the average values of at least three tests were taken.

The measurements were recorded continuously, and the friction force was recorded with the load cell. Temperature of samples was measured with a thermocouple connected to the follower. Tribological properties were determined at 40℃, in 70% relative humidity (RH). Both boron compound additives used were 99.9% pure, and the size of HBN and BA powders was as of 10 and 50 micrometers, respectively. The results were evaluated taking into consideration the forming a tribofilm in mixed or EHD lubrication.

IV. Results and Discussion

A. Tribological Behavior

Fig. 3 shows normal force, friction force and friction coefficient in a cam period lubricated by base oil ((20W50 engine oil) at 312 N, 600 rpm, 40°C. It is seen that the minimum contact load between the cam follower is occurred at the cam base circle, called preload value, the minimum contact load increases along with the cam profile and reaches its maximum value at the cam nose i.e. at cam angles around degrees 90.

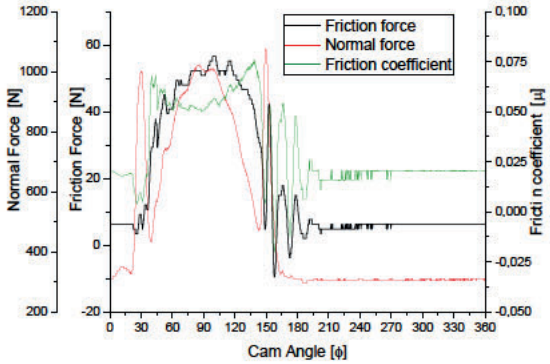


Figure 3: Normal force, friction force and friction coefficient in a cam period lubricated with the base oil (20W50 engine oil) at 312 N, 600 rpm and °C 40.

Figure 4 shows the friction force as a function of contact speeds to be able to diagnose the effect of the hBN additives on the base engine oil (20W50) at 400, 600 and 800 rpm. As shown in the Fig. 4, the friction force of the base oil has decreased with the addition of hBN. However, there is a slight difference between the friction force of the base oil and boron compounds of hBN. The friction force of base oil and +4 wt% hBN have a different value. In the tests performed at 400, 600 and 800 cycles, it was observed that the percentage of reduction in friction forces caused by the addition of 4%hBN additives to the base oil was 3.5%, 5.7% and 4%, respectively. In order to determine the reduction rates in friction values more precisely, the Distance Weighted method has been applied to the curves in the graphs.

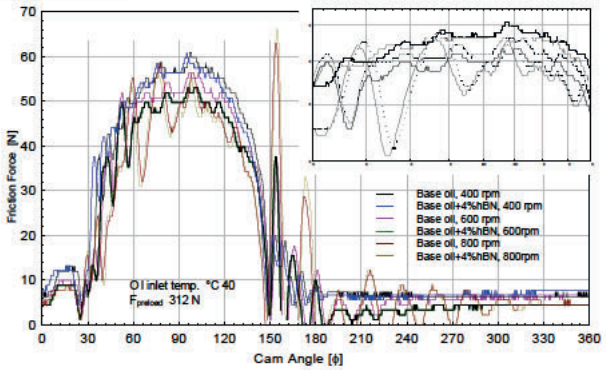


Figure 4: Test results for friction forces at 400, 600 and 800 rpm with the base oil and 4% hBN.

Figure 5 shows the friction force as a function of contact speeds to be able to determine the effect of the BA additive on the base engine oil at 400, 600 and 800 rpm. As shown in Fig. 5, the friction force of the base oil was decreased with the addition of BA boron compound. However, there is a more difference between the friction force of the base oil and BA compound than that of hBN. Indeed, at 400, 600 and 800 cycles, it was observed that the percentage of reduction in friction forces caused by the addition of 4%BA additive to the base oil was 10%, 15% and 11%, respectively.

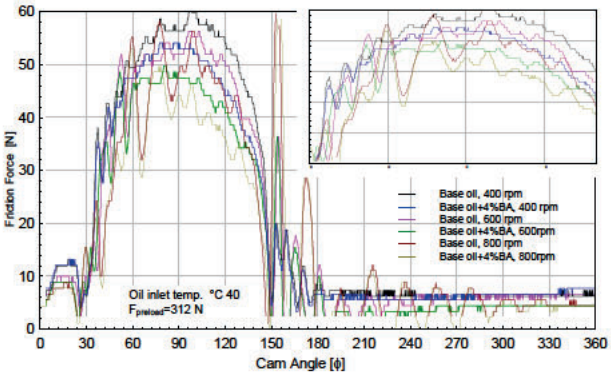


Figure 5: Test results for friction forces at 400, 600 and 800 rpm with the base oil and 4% BA.

Furthermore, in order to investigate the effect of contact load on friction force in base oil, 4% hBN and 4% BA, the tests performed at two preload values, ($F_{preload}$=212N and 412N). As shown in Fig. 6, it was observed that, reducing effect on the friction, addition of 4% hBN and 4% BA compounds to the base oil, was 7% and 12% at 212 N preload, and 7.7% and 13.2% at 412 N preload, respectively. These results indicate that the friction reduction effects of boron additives on the contact loads exist, and these additives improved the antiwear property of base engine oil. The results obtained are consistent with those of Shah [5] and Baldwin [6]. It's clearly seen that the boric acid additive has more influence than that of hexagonal boron nitride additive in reducing friction force, and this influence increased with increasing contact load at 40°C, as shown in Fig. 6.

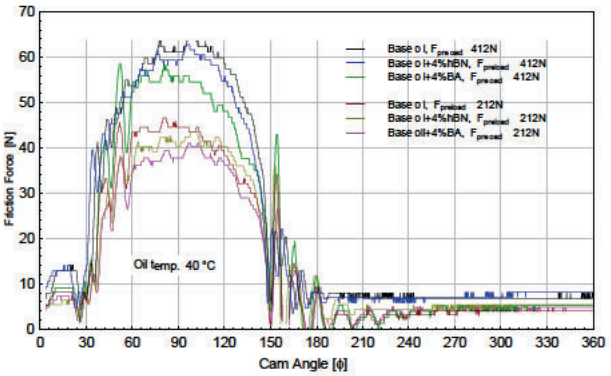


Figure 6: Effect of load on friction force in base oil, 4% hBN and 4% BA tests.

V. Conclusion

Contact loads of the cam and follower were 212, 312 and 412 N, respectively. These were very near Hertzian pressure becomes in EHL, and contact surface non-conformal. Therefore, the system may be evaluated in a mixed lubrication regime, and local micro EHL zones may form at the interface of roughness surfaces. Local micro EHL zones and micro contact of the roughness surfaces affects the friction coefficient seriously. Especially, the formation of tribofilm in micro scale can reduce friction coefficient in the boundary or mixed lubrication. Furthermore, the tribofilm reduces the effect of micro EHL and micro contact of the surface roughness. From the above experimental results, following conclusions can be drawn:

- In engine parts, especially on the cam and follower, boron compound additives can reduce friction force seriously, forming a tribofilm, in the boundary or mixed lubrication conditions.
- The reduction of friction force, especially at high speeds, and high loads was observed.
- The effect of boron compound additives on the friction force rises with increasing of contact load and speeds.
- Effect on friction force of BA and hBN on the pair of contact was very close to each other. But the effect of boric acid additive was better than that of hBN.
- Boric acid additives have more influence than that of hexagonal boron nitride additive in reducing friction force, and this influence was increased with increasing contact load.

Acknowledgment

This research was funded by the Karadeniz Technical University, Scientific Research Projects Unit, (Project Code: 2008.112.003.9). Also, author would like to thanks to National Boron Research Institute for supporting of boron compounds.

References

[1] H. Hirani, Mixed Lubrication. In *Fundamentals of Engineering Tribology with Applications* (pp. 326-370). Cambridge: Cambridge University Press. 2016.

[2] H. Bas and Y. E. Karabacak, “Investigation of the Effects of Boron Additives on the Performance of Engine Oil,” Tribology Transactions, 57:4, 740-748, 2014.

[3] T. Bartels, “Lubricants, 5. Gear Lubrication Oils”, Ullmann's Encyclopedia of Industrial Chemistry, (2011), ISBN: 9783527306732.

[4] J. M. Herdan, “Friction modifiers in engine and gear oils”, Lubrication. Science, 12, 3, pp. 265-276, 2000.

[5] F. U. Shah, “Boron Compounds as Additives to Lubricants, Synthesis, Characterization and Tribological, Optimization”, Lulea University of technology, 2009.

[6] B. A. Baldwin, “Relative antiwear efficiency of boron and sulphur surface species”, Wear, 45, 3, pp. 345-353, (1977).

[7] Y. Kimura, T. Wakabayashi, K. Okada, T. Wada, and H. Nishikawa, “Boron nitride as a lubricant additive”, Wear, 232, 199, pp. 206.56, 1999.

[8] A. Erdemir, “Advances in Boron-Based Lubricants and Lubrication Additives”, IV International Boron Symposium, October Eskisehir, TURKEY, 2009.

[9] F. Chinas-Castillo and H. A. Spikes, “Mechanism of action of colloidal solid dispersions”, J. Tribology-Trans., ASME, 125, 3 pp. 552-557, 2003.

[10] P. J. Pacholke and K. M. Marshek, “Improved worm gear performance with colloidal molybdenum-disulfide containing lubricants”, Lubr. Eng., 43, pp. 623-628, 1987

[11] R. S. Barnett, “Molybdenum disulfide as an additive for lubricating greases”, Lubr. Eng., 33, 308-313, 1977.

[12] A. Erdemir, “Solid lubricants and self-lubricating films, in “Handbook of Modern Tribology,” B. Bhushan, ed , CRC Press, Boco Raton, FL, pp. 787-818, 2001.

[13] G.W Rowe, “Some observations on the frictional behaviour of boron nitride and of graphite”, Wear, 3, 4, pp. 274–285, 1960.

[14] J. M. Martin, T. LeMogne, C. Chassagnette, and M. N. Gardos, “Friction of hexagonal boron nitride in various environments”, Tribol. Trans., 35, 463-472, 1992.

[15] C. C. Klepper, J. M. Williams, J. J. J. Truhan, J. Qu, L. Riester, R. C. Hazelton, J. J Moschella, P. J. Blau, J. P. Anderson, O. O. Popoola and M. D. Keitz ”Tribomechanical properties of thin boron coatings deposited on polished cobalt alloy surfaces for orthopedic applications”, Thin Solid Films, 516, 10, pp. 3070–3080, 2008.

[16] M. Lovell, C. F. Higgs, P. Deshmukh and A. Moble, “Increasing formability in sheet metal stamping operations using environmentally friendly lubricants”, J. Mat. Pro.Tech. 177, pp. 87-90, 2006.

[17] J. H. Wu, B. S. Phillips, W. Jiang, J. H. Sanders, J. S. Zabinski, and A. P. Malshe, “Bio-inspired surface engineering and tribology of MoS2 overcoated cBN–TiN composite coating”, Wear, 261, pp. 592–599, 2006.

[18] H. U. Jamali, A. Al-Hamood, O.I. Abdullah, A. Senatore and J. Schlattmann, “Lubrication Analyses of Cam and Flat-Faced Follower,” Lubricant, 7, 31, 2019

[19] C. M. Taylor, “Valve train lubrication analysis. Seventeenth Leeds-Lyon Symposium on Tribology—vehicle tribology, Tribology Series 18, , pp. 119–131 (Elsevier, Oxford). 1991.

[20] H. A. Spikes, V. Anghel and R. Glovnea, “Measurement of the Rheology of Lubricant Films Within Elastohydrodynamic Contacts” Tribology Letters, 2014

[21] H. Bas, Ph.D. Thesis, ‘‘Investigation of Friction Behavior of Additive Engine Oils on Cam Mechanisms,”Karadeniz Technical University Institute of Natural and Applied Sciences, 2001.

Investigation of Effects of Surface Roughness on the Performance of Cam Mechanisms

Hasan BAŞ[1]

[1] Karadeniz Technical University, Trabzon/Turkey, bas@ktu.edu.tr

Abstract **- The surface roughness of two contacting surfaces significantly influences the tribological performance of the mechanical elements. Their impression is more pronounced under the mixed elasto-hydrodynamic lubrication condition. The cam and flat follower mechanism is a typical sample in which adverse tribological conditions, including direct boundary interactions occurs. In this study, effects of surface roughness on the friction force and friction coefficient are investigated using engine oil at different test conditions in a cam follower mechanism. It is seen that decreasing roughness of the contact surfaces has a more desirable tribological performance, and decrease friction coefficient, therefore increase wear resistance.**

Keywords **- Surface Roughness, Oil Film, Friction Coefficient, Combustion Engine Oil**

I. Introduction

The effect of surface roughness pattern on the tribological performance of mechanical elements is an important issue that has attracted much attention for more than four decades. The properties of surface layers are crucial to understand the tribological performance. Various machine components such as bearings, gearbox, camshaft, crank shaft, etc. require excellent tribological properties and practices to ensure longer life. Understanding engineered surfaces is very important for solving many scientific problems that involve friction, contact mechanics, heat conduction, electric current conduction, and component design [1]. Surface roughness has been the subject of experimental and theoretical investigation for a long time. Surface texture is one of the most important parameters that controls the tribological behavior. Patir and Cheng developed a stochastic analysis approach to study the effects of surface roughness height and distribution between components upon the lubrication of machine elements, [2], [3]. The approach has also been used to assess cam and follower performance e.g. Refs. [4], [5]. For the direct acting cam and follower situation shown in Fig. 1, Dowson et al. [3] studied the predicted effects of surface roughness upon nominal film thickness, power loss and load carried by the asperities. Due to the fact that there are still very few examples of attempting to reduce friction in boundary lubrication conditions with high contact pressure such as cam/follower combination, the effects of surface texturing remain unknown. Ehret et al. [6] have studied the effects of various surface textures, described by either an orientated waviness or a uniform distribution of asperities, for rolling and sliding conditions. Surface texture can lead to important differences and in pure sliding it is shown that the best lubrication condition is produced for transverse waviness. The definitive linking of surface texture to improved machine element performance in, say, a cam and follower, would be a powerful technique [5]. O'Connor and Spedding [7] undertook an experimental study of the effect of surface waviness on a cam profile. Such waviness was influenced by the form imposed onto the grinding wheel by the roller dresser. Wear and oil retention characteristics of the automobile camshaft surface were investigated. It was suggested that the wavy cam surface could offer good oil retention characteristics but that with time the release of hard wear particles due to the wave peaks could be problematical. It was proposed that a cam surface having the same roughness, but zero or reduced waviness, would show improved performance. Waviness on automotive cams remains an issue and it is interesting to note that analysts are now in a position to contribute to the debate on the overall influence [5], [8]. For tribologists, the most critical interface in the valve train is the cam and follower due to the problem of providing effective lubrication. Often, it was assumed that the cam and follower operated in the boundary lubrication regime; however, more recent investigations have shown that mixed and elastohydrodynamic lubrications are the dominant regime.

The cam and follower lubrication is one of the interesting issues of transient EHDlubrication that has rarely been investigated specially for rough surfaces. Thus, effects of surface roughness on the friction force and friction coefficient are investigated. The results can provide guidance on the tribological application of cam and follower in industrial application.

II. Theoretical Bases

A. Force Analysis

In the cam mechanism, the contact load acting on the cam-follower contact area varies depending on the cam rotation angle. Maximum contact load occurs at cam angles where the cam nose is in contact with the follower.

The loading state of the roughness flat-surface follower-cam pair is shown in Figure 1. Here, the contact load is in the same direction as the centerline during a cam period due to the flatness of the follower surface. The contact load at the contact surface is constant only along the base circle. The cam surface acts on the follower surface with the cam pressure angle depending on the cam rotation angle. The maximum value of the pressure angle in question occurs in the cam position

where the contact point of the two surfaces is furthest from the center line, as shown in Fig.1. Thus, the cam remains under the influence of the maximum torque at the rotation angles corresponding to these positions.

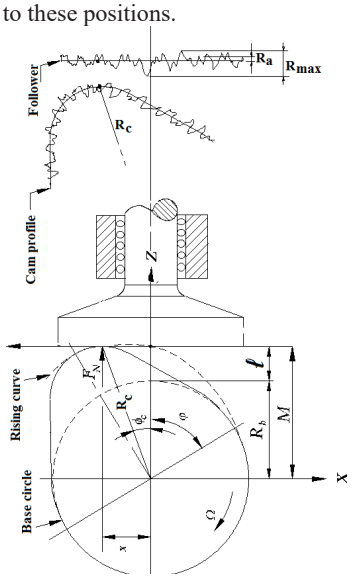


Figure 1. Roughness considerations in a cam and flat surfaced follower pair.

B. Friction Force

The general form of the Reynolds equation for EHD lubrication of two smooth surfaces has been presented in many publications, such as [9]. Generally, when two surfaces contact each other in the presence of lubricant, pressure in the entire domain is governed by:

$$\frac{\partial}{\partial x}\left(\frac{\rho h^3}{\eta}\frac{\partial p}{\partial x}\right)+\frac{\partial}{\partial y}\left(\frac{\rho h^3}{\eta}\frac{\partial p}{\partial y}\right)=6u\frac{\partial}{\partial x}(\rho h)+12\frac{\partial \rho h}{\partial t} \quad (1)$$

where p is the hydrodynamic pressure, ρ is the lubricant's density, and η is the viscosity of the lubricant. Film thickness, h, is defined as the gap between two surfaces. Therefore, when the surfaces are rough, the film thickness for the line contact problem can be expressed as following:

$$h(x,y)=h_0+v(x,y)+\delta(x,y) \quad (2)$$

where V and δ are the surface elastic deformation and the asperities' height, respectively. The regime of true mixed lubrication, where there are both elâstohydrodynamic and boundary lubricated regions within a single contact, remains quite poorly understood. The cam-follower pair is a very complex lubricated contact because its continuous variation of load, speed and radius of curvature. In order to determine the friction force arising in the elasto-hydrodynamic contact under the forces acting through a period in the cam mechanism, instead of h mentioned above, using which is the film thickness at the lubricated maximum hertz contact between cam and follower, the final calculation process can be write as follows.

$$F_s=\frac{\eta_0(u_1-u_2)}{h_{cen}}\cdot\ell\cdot\int_{-b}^{b}e^{\alpha\cdot p_{\max}\sqrt{\left(1-\frac{x^2}{b^2}\right)}}dx \quad (3)$$

III. Experimental Setup

A. Description of The Test Rig

The test apparatus was manufactured in the Mechanical Engineering Department at Karadeniz Technical University. Detailed information about the test apparatus can be obtained from Ref. [10], [11], [12]. The mechanical system and measurement unit of the test apparatus are designed to measure the frictional behavior between cam and follower at different loads, speeds and lubricating conditions. A lubricating system is designed to investigate the effect of different type of oils on the friction behavior at the cam / follower interface. In order to ensure safe and accurate data collection and storage, a computer and a data acquisition system are attached to the measurement unit and a computer program is written to evaluate this data.

B. Measurement Procedures

In the test apparatus, two types of accelerometers were used to measure the mass force in vertical and horizontal directions due to acceleration of moving masses. The mass force in the horizontal direction is measured by using a piezoelectric type of accelerometer sensor, which works by a compression effect and is mounted rigidly between two linear bearings in the loading system. To determine the friction force and friction coefficient, detailed information about friction force applied to the follower by cam, causes a tensile effect on the measurement beam and the effect produces a signal proportional to the beam force, which can be used in the calculation of the friction force [12]. The measurements were recorded continuously, and friction force was recorded with the load cell. Temperature of samples was measured with a thermocouple connected to the follower. Tribological properties were determined at ambient temperature, in 70% relative humidity (RH). The results were evaluated taking into consideration the forming a tribofilm in mixed or EHD lubrication.

In order to understand the effects of surfaces roughness on the tribological behavior in the cam-follower mechanism, surface roughness and morphologies of the contact surfaces and wear surfaces were investigated using a surface roughness test apparatus and scanning electron microscopy. In this study, roughness of the surfaces of cam and followers are measured using contact profilometers. The most common method of studying surface texture features is the contact-type stylus profilometer. A fine, very lightly loaded stylus is dragged smoothly at a constant speed across the surface under examination. The height position of the stylus generates an analog signal which is converted into a digital signal that is stored, analyzed, and displayed, and the output provides a magnified view of the original profile.

Table 1. Chemical composition of the cam and follower.

	C	Si	Mn	P	S	Cr	Ni	Mo	Cu	Al	Ti
Cam	3.42	2.15	0.6	0.05	0.01	0.9	0.95	0.05	0.10	0.01	0.004
Follower	1.06	0.3	0.3	0.007	0.01	1.48	0.13	0.04	0.23	0.013	0.003

A photograph of 2D contact-type profilometer and the profile of mean roughness (Ra), roughness depth (Rz) and maximum roughness (Rmax) formulations are shown in Fig. 2.

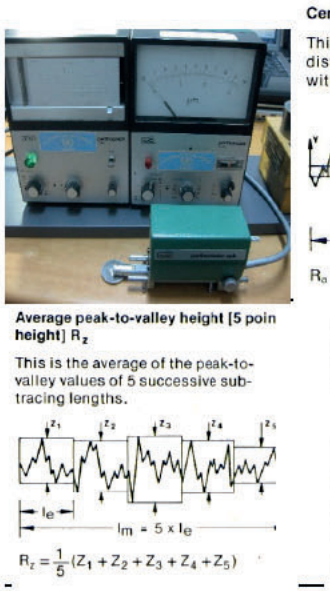
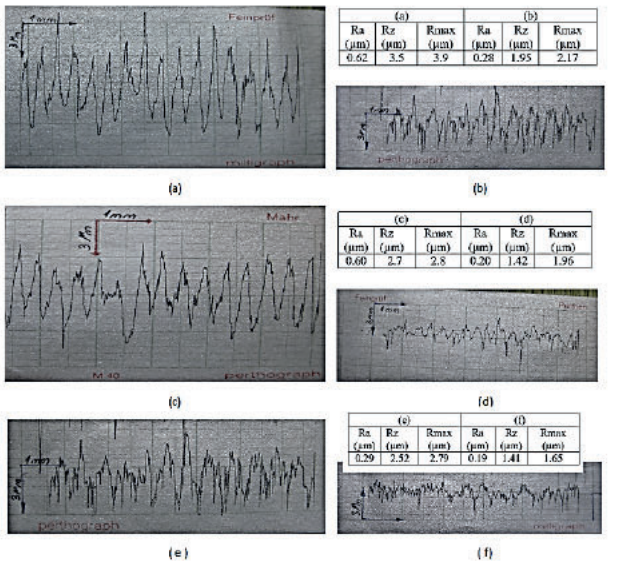


Fig. 2. Photograph of 2D contact-type stylus profilometer and roughness formulations of Ra, Rz and Rmax,



	(a)			(b)		
	Ra (μm)	Rz (μm)	Rmax (μm)	Ra (μm)	Rz (μm)	Rmax (μm)
	0.62	3.5	3.9	0.28	1.95	2.17

	(c)			(d)		
	Ra (μm)	Rz (μm)	Rmax (μm)	Ra (μm)	Rz (μm)	Rmax (μm)
	0.60	2.7	2.8	0.20	1.42	1.96

	(e)			(f)		
	Ra (μm)	Rz (μm)	Rmax (μm)	Ra (μm)	Rz (μm)	Rmax (μm)
	0.29	2.52	2.79	0.19	1.41	1.65

Fig.3. Graphical representation of the measured roughness profile of the cam and followers, (a) before, (b) after non-lubricated (dry) tests of follower1; (c) before, (d) after lubricated tests of follower 3; (e) before, (f) after lubricated test of the cam.

Fig. 3 Shows, roughness measurements of the worn surfaces of cam and follower before and after lubricated and non-lubricated tests were performed using a surface roughness tester (Mahr-Germany Perthen Perthometer) pursuant to DIN 4768, as a measure of (Ra), (Rz) and (Rmax).

It is seen that the result of surface roughness of the cam and follower tests with non-lubricated conditions are higher than lubricated conditions.

C. *Test Methodology*

In this study, to investigate tribological behavior of surface roughness on the cam mechanism, a new automotive cam shaft and four followers used. Followers has a different roughness manufactured by grinding from the same material. Chemical composition of the cam and follower were given in Table 1. In the manufactured test apparatus, cam is direct acting to the follower. Followers with different surface roughness initially run against nominally same cam for at a fixed camshaft speed and temperature. Experiments were conducted at 600 rpm, and 312 N contact load. Experimental periods were selected as 10 minutes for every test after running-in, and the average values of at least three tests were taken.

IV. Results and Discussion

A. *Tribological Behavior*

Fig. 4 shows normal force, friction force and friction coefficient in a cam period lubricated by base oil (20W50 engine oil) and non-lubricated i.e. dry condition at 312 N, 600 rpm, 40°C. It is seen that, in case of lubricated and dry condition, the maximum contact load increases along with the cam profile and reaches its maximum value at cam nose, (i.e. at cam angles around 90^0). Both of lubricated and non-lubricated situation, the value of contact load around the cam nose has same value at the same preload. However, the fric-

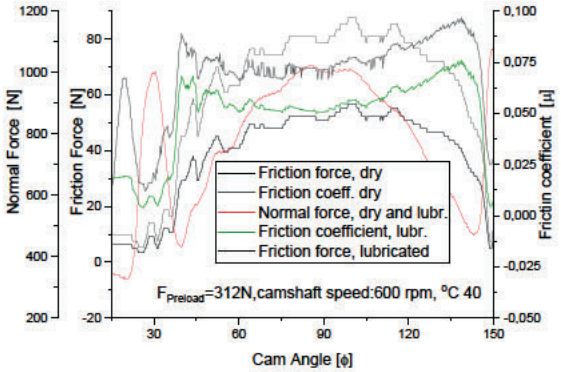


Figure 4: Normal force, friction force and friction coefficient in a cam period lubricated with the base oil (20W50 engine oil) at 312 N, 600 rpm

tion force and thus the friction coefficient are different. It is seen that the friction values in the oil-free contact are greater than that lubricated contact.

In order to investigate the tribological behavior of surface roughness in dry and lubricated contact, two followers were used with very close surface roughness Ra value to each other. Figure 5 illustrates the variation of the surface roughness with the number of cycles when used two followers under dry and lubricated conditions. Both of roughness values of the followers are decreases with increasing cycle number of the cam. It is seen that the slope in the curves approaches zero after 3000 turns. Based on these results, it can be stated that the oily contact has a rather abrasion effect and is subject to abrasion, whereas in dry contact, the adhesion effect is more intense due to scuffing on the surface. In the running-in period, the asperities are gradually flattened and friction coefficient is progressively reduced until reaching the steady state.

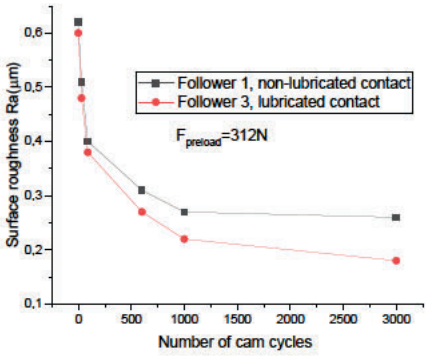


Figure 5. Variation of friction coefficient with number of cycle using follower1 and follower2 under dry and lubricated conditions.

It is important to note that the surface texture of both followers (different surface roughness values) plays an important role in controlling the frictional behavior. Figure 6 illustrates the variation of coefficient of friction with surface roughness under dry and lubricated conditions. The friction coefficient as a function of contact roughness to be able to diagnose the effect of roughness on the lubricated and non-lubricated condition were done. Surface roughness of follower1 and follower3 are given in Fig.3 above. Ra, Rz and Rmax values of follower2 and 4 are 0.16[μm], 1.18[μm], 1.22[μm] and 0.13[μm], 0.52[μm], 0.58[μm], respectively. As shown in Fig. 6, the friction coefficient of the follower1 is slightly higher than that of the follower2 under the non-lubricated condition. In addition, the friction coefficient values of the follower3 is quite higher than that of follower4 under the lubricated condition. All friction coefficient curves showed the same trend. It was observed that the percentage of reduction in friction coefficient under dry conditions were higher than that of lubricated condition. Hence, it is understood that in dry contact, the adhesion effect is more intense due to scuffing on the cam-follower surfaces.

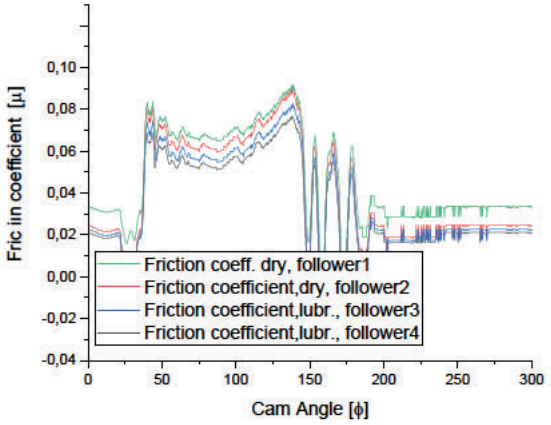


Figure 6. Variation of friction coefficient with cam angle when used follower3 and follower4 under dry and lubricated conditions.

To identify the effect of roughness on lubricated and non-lubricated state, frictional force tests were made as a function of contact roughness. Figure 7 shows the variation of friction force with surface roughness under bot of dry and lubricated conditions. As shown in Fig. 7, the friction coefficion of the follower1 is slightly higher than that of the follower2 under the non-lubricated condition.

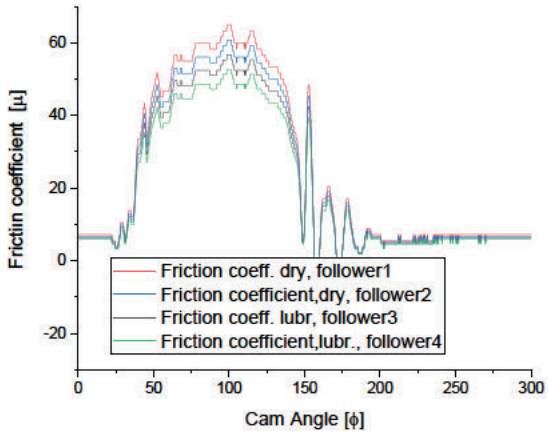


Figure 7. Variation of friction force with cam angle when used follower3 and follower4 under dry and lubricated conditions.

And also, the friction coeffiicion values of the follower3 is quiet higher than that of follower4 under the lubricated condition. All friction coefficient curves showed the same trend. It was observed that the percentage of reduction in friction coefficient under dry conditions were higher than that of lubricated condition.

According to the above results, it is clear that the surface roughness indeed has an important role on of friction values during sliding. Thus, surface textures were characterized by means of surface roughness parameters.

Results displayed in Figure 6 and Figure 7 shows that for rising values of roughness is increasing, friction coefficient and friction force increases too. It can be explained that high roughness leads to smaller contact areas leading directly to

high local contact pressures on the contrary to profiles with a low Ra.

Figures 8a–8c shows scanning electron microscope morphologies of the contact surfaces before test, and Figures 8d–8f after dry test under 312 preloads at 600 rpm and ambient temperature.

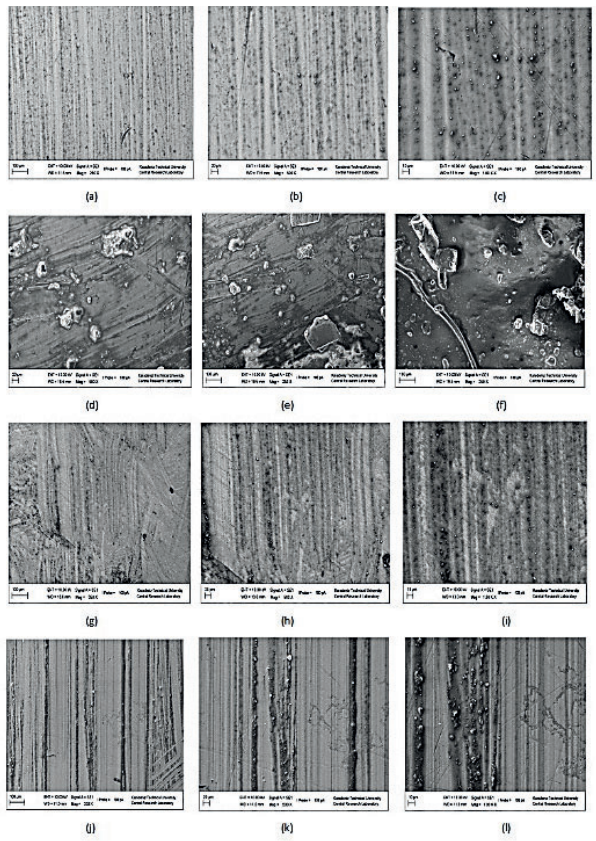


Fig. 8—SEM morphologies of contact surfaces; non-lubricated i.e. dry contact tests (a)-(c) before test and (d)-(f) after test; lubricated contact tests (g)-(i) before test and (j)-(l) after test under 312 preload, at 600 rpm. In this chapter, attempts have been made to study various surface roughness values.

As seen in Figs. 8d–8f, the worn surface of the follower lubricated by base oil was characterized by slight adhesion wear. For lubricated conditions, Figs. 8g–8i and Figs. 8j–8l shows morphologies of the contact surfaces before and after the tests, respectively. As seen in Figs. 8j–8l, the worn surface of the follower was characterized by abrasion wear, some adhesion wear.

V. Conclusion

Contact loads of the cam and follower were 312 N. These were very near Hertzian pressure becomes in EHL, and contact surface non-conformal. Therefore, the system may be evaluated in a mixed lubrication regime, and local micro EHL zones may form at the interface of roughness surfaces. From the above experimental results, following conclusions can be drawn:

- the surface roughness has an important role on of friction values in engine parts during sliding, especially on the cam and follower,
- in dry contact, the adhesion effect is more intense due to scuffing on the cam-follower surfaces.
- In the running-in period, the asperities are gradually flattened and friction coefficient is progressively reduced until reaching the steady state.
- as the roughness values increase, the friction coefficient and friction force also increase.

Acknowledgment

This research was funded by the Karadeniz Technical University, Scientific Research Projects Unit, (Project Code: 2008.112.003.9

References

[1] A. Torabi, S. Akbarzadeh, M. R. Salimpour and M. Taei, “Effect of surface roughness pattern on transient mixed elastohydrodynamic lubrication” Surface Topography: Metrology and Properties, Volume 4, Number 1, 2015

[2] N. Patir, H.S. Cheng, “An average flow model for determining effects of three-dimensional roughness in partial hydrodynamic lubrication” J. Lubr. Technol., Trans. ASME 100 1 1978 12–17.

[3] D. Dowson, C.M. Taylor, G. Zhu, Mixed lubrication of a cam and flat faced follower, Proc. 13th Leeds–Lyon Symposium on Tribology — Fluid Film Lubrication, Tribol. Ser. vol. 11 Elsevier, Amsterdam, 1987, pp. 599–609.

[4] L.S. Yang, A. Ibo, H. Negishi, A valve train friction and lubrication analysis model and its application in a camrtappet wear study, SAE Technical Paper 962030, 1996, 9 pp.

[5] [5] M. Priest, C.M. Taylor “Automobile engine tribology — approaching the surface” Wear 241 2000 193–203.

[6] P. Ehret, D. Dowson, C.M. Taylor, Time-dependent solutions with waviness and asperities in EHL point contacts, Proc. 23rd Leeds–Lyon Symposium on Tribology — Elastohydrodynamics ’96, Tribol. Ser. vol. 32 Elsevier, Amsterdam, 1997, pp. 313–324.

[7] R.F. O’Connor, T.A. Spedding, Investigation into the effects of waviness present on the cam surface profile of an engine camshaft, Precis. Eng. 11 2 1989 83–88, Butterworth.

[8] I. Tanimoto, M. Kano, M. Sasaki, Establishment of a method for predicting cam follower wear in the material development process, SAE Technical Paper 902087, 1990, 8 pp.

[9] Zhu D and Wang Q J 2013 Effect of roughness orientation on the elastohydrodynamic lubrication film thickness ASME J. Tribol. 135 031501

[10] H. Bas, Ph.D. Thesis, ‘‘Investigation of Friction Behavior of Additive Engine Oils on Cam Mechanisms,’’ (in Turkish) Karadeniz Technical University Institute of Natural and Applied Sciences, 2001.

[11] H. Bas, A. Bıyıklıoglu, ‘‘Investigation of Frictional Behavior of Different Kinds of Engine Oils on Cam Mechanisms Used in Automotive Engines,’’ International Conference on Lubrication Techniques, Istanbul, 87–103, ICOLT’99.

[12] H. Bas, A. Bıyıklıoglu and H. Cuvalci, “A New Test Apparatus For the Tribological Behavior Of Cam Mechanisms,” Experimental Techniques, 2003.

Investigation on Effect of Shrinkage Allowance to the Fatigue Life of Compound Cylinders Operating at High Pressure

M.AYDIN[1] Y.F.YAPAN[2] and M.TURKOZ[1]

[1] Konya Technical University, Konya/Turkey, maydin@ktun.edu.tr
[1]Konya Technical University, Konya/Turkey, mturkoz@ktun.edu.tr
[2]Turkav R&D, Konya/Turkey, yusuffurkanyapan@gmail.com

***Abstract* - In this study, the fatigue life of a shrink-fitted two-layer cylinder operating under high pressure is estimated by using Finite Element Method (FEM). The cylinders material used in this investigation is AISI 4140 steel. In this context, the fatigue life of the inner and outer cylinders was determined under 300 MPa pressure depending on the one-sided shrinkage tolerance. In the study, 11 different shrinkage allowance were used. The optimum shrinkage allowance was determined according to the fatigue life of the two cylinders. Accordingly, when the shrinkage allowance increased, the life of the inner cylinder increased up to 189.73 times, while the life of the outer cylinder decreased by 99.8%. In addition, the fatigue life values of the two-layer cylinders were compared with the fatigue life of the single-layer cylinder with the same geometric dimensions. While the life value was obtained as 5586 in the single layer cylinder, this value was calculated as 9530 even in the double layer cylinder with a shrinkage allowance of 0.01 mm. As a result, it has been found that using shrink-fitted two-layer cylinders are more advantageous in terms of fatigue life than single-layer cylinders.**

***Keywords* – Fatigue Life, Finite Element Method, Shrinkage Allowance, Two-Layer Cylinder**

I. INTRODUCTION

Shrink fitting of cylinders is a widely used assembly method in the industry. For example, gear-shaft connections, pressure intensifier cylinders, tool holders, turbine shafts and rotor-electric motor connections [1]. Shrink fit is an assembly method in which heat is used, high radial stresses occur in cylindrical parts, and two or more cylinders with a tolerances are combined. [2]. The heated metal cylinder will expand and grow, and the cold cylinder can be easily assembled inside the heated cylinder. Then the heated cylinder, which has cooled down, tries to return to its original size and applies radial pressure to the inner cylinder. [1]. Friction between the cylinders is very important in terms of torque transmission and safety of the assembly.

Öztürk [1] investigated analytically and numerically the interference pressure of two cylinders designed with steel-steel and steel-aluminum materials at different shrinkage allowance in the shrink fit process. In addition, the stresses occurring along the thickness distribution in the shrink fit process were determined and verified with analytical values. In the study, it was stated that when the geometry of cylinders is simple, it may be sufficient to calculate the stress and interference pressures by analytical methods, but the finite element analysis results are more reliable in complex geometry assembly processes. Patil A. S., [3] used cylinders made of AISI 1020 steel with shrink fitted inner diameter 100 mm and outer diameter 153 mm. The cylinders were subjected to cyclic internal pressure between 0-250 MPa and the constant pressure applied to the outer surface. Both the cylinders were subjected to constant and non-constant amplitude loading. Fatigue analysis of compound cylinders was done by FEM. Due to the high number of cycles, the Stress-Life method was used. Analysis results showed that the outer cylinder was damaged earlier than the inner cylinder. It was stated that the contact surfaces, which are the most important region in the cylinders, fatigue faster than the other regions. Özel et al., [4] investigated the stress and deformations for different hub-shaft connections using FEM to determine the shrink fit between compound cylinders. It was stated that the ends of the cylinders are more dangerous than the center of cylinders. Miraje et al. [5] analyzed the fatigue life of three compound cylinders under internal pressure was determined by Finite Element Analysis. In addition, three cylinders were designed to operate at infinite life, and the total cylinder volume was tried to be reduced to a minimum. Optimized parameters in the study were cylinder diameter ratios and shrinkage allowance. The optimum limits of the geometric parameters of all three cylinders of limited life were determined and verified through analytical and analysis.

High pressures are used in the industry in hydroforming, cutting with water jet, in the strength and fatigue tests of technological products such as cylinders, pipes, valves, fittings, hydrogen tanks, especially composite tanks. In addition, ultra-high pressures are also used in autofrettage, industrial cleaning and high-pressure pasteurization processes. Single or multi-layered cylinders are used in the high pressure regions of the pumps called pressure intensifier, which are used to getting high pressures.

In this study, the effect of shrinkage allowance on fatigue life in compound cylinders was investigated. In the literature, it is

known that the radial stresses during occurring the shrink fitting process both increase the fatigue life of the material and reduce the total stress on the material [6-10]. The optimum shrinkage allowance is important on fatigue life of cylinders under high pressure. In this analysis study, the optimum shrinkage allowance of the two-layer AISI 4140 steel working at high internal pressures was determined by FEM analysis.

II. Materials and Methods

In this study, the effect of shrink fit tolerances on fatigue life of 2-layer shrink-fit cylinders with different tolerances operating under cyclic 300 MPa high pressure was investigated. For this, the fatigue tests of the process were carried out via Finite Element Analysis (FEA).

A. Geometric Modelling

In the process, two cylinders were used as inner and outer (Figure 1). The geometric dimensions of the cylinders are given in Table 1. Both cylinders are 30 mm thickness and 50 mm length.

Table 1: Geometric Dimensions

Part	Inner Diameter	Outer Diameter
Inner Cylinder	50 mm	(80+δ) mm
Outer Cylinder	80 mm	110 mm

The inner and outer cylinders in the process were created in 3D in the Solidworks program. The shrinkage allowance is given in the FEA at the stage of defining the contact conditions. These two cylinders were assembled as layers and transferred to ANSYS 2020 Student Version Finite Element Software as a parasolid file.

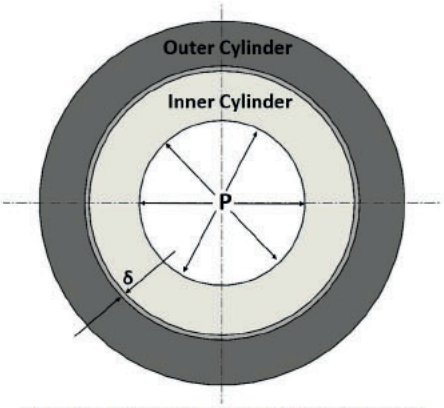


Figure 1: Schematic representation of process

B. Defining Material Properties

Tempered at 540°C & oil quenched AISI 4140 steel material is used in both cylinders. The mechanical properties of the material are defined in the ANSYS material library (Table 2).

Table 2: Mechanical properties of the material [11]

Yield Strength	Tensile Strength	Elastic Modulus	Poisson Ratio
835 MPa	952.5 MPa	210 GPa	0.29

In order to perform fatigue analysis, the S-N diagram of the material is defined in the software for different stress ratios (R-Ratio) values. The S-N curves given in Figure 2 are taken from the CES Selector 2016 program. The fatigue life scale is shown logarithmically.

C. Contact Definitions

Shrinkage allowance from 0.01mm to 0.11mm was defined at this stage. A friction coefficient of 0.2 was applied between the two cylinders.

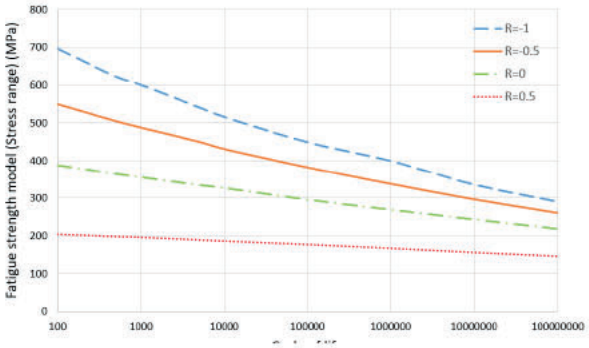


Figure 2: S-N curves for different R-Ratio values of the material [11]

D. Meshing

For both cylinders, 'Hexa' mesh type of 5 mm size was used in the 'Multi Zone' method. The model has 3600 elements and 19320 Nodes in total.

E. Defining Boundary Conditions

Two boundary conditions were defined in the SEA model. As can be seen in Figure 3, the boundary condition named A shown in red represents the internal pressure of 300 MPa applied inside the cylinders. With the boundary condition named B shown in blue, the cylinder is allowed to expand in axial and radial directions, while its rotation in tangential direction was limited.

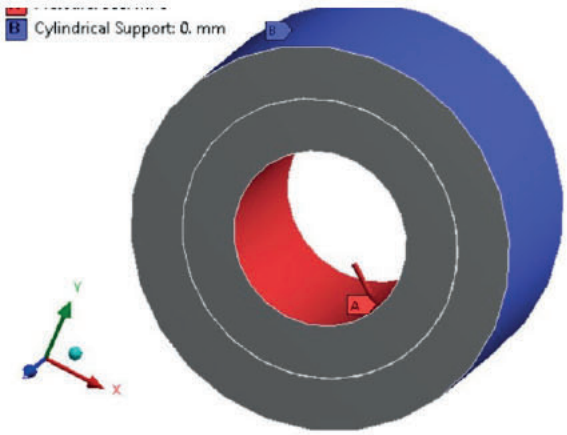


Figure 3: Boundary conditions

F. Fatigue Analysis

In order to perform the fatigue analysis, first of all, static structural analysis should be performed. Because of the shrinkage allowance, a pre-stress occurred in the compression direction in the inner cylinder and in the tensile direction in the outer cylinder. Due to the internal pressure, tension occurred in both cylinders in the tensile direction (Figure 4).

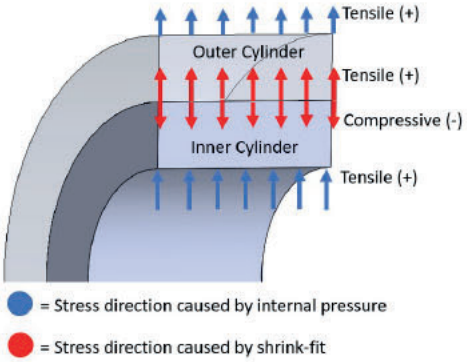


Figure 4: Stress directions

First, the Von-Mises equivalent stresses resulting from the shrink-fit without applying pressure inside the compound cylinders were determined. The value obtained is the minimum stress (S_{min}). Then, by applying an internal pressure of 300 MPa, Von-Mises equivalent stresses were determined. The value obtained here is the maximum stress (S_{max}). The stress ratio given in Equation 1 is the ratio of the smallest stress occurring in the structure to the largest stress during loading.

$$R = S_{min} / S_{max} \quad (1)$$

For each shrinkage allowance, R-Ratio values occurring in the inner and outer cylinders were determined. The compound cylinders are proportionally loaded at constant amplitude. Figure 5 shows S_{max}, S_{min} and R-Ratio value determined in the inner cylinder for 0.01 mm shrinkage allowance. In ANSYS fatigue tool, the loading type was selected as "Ratio" and the loading ratio was entered into the software as -0.05098. Besides the R-Ratio value in the outer cylinder was determined and entered in a second ANSYS fatigue tool.

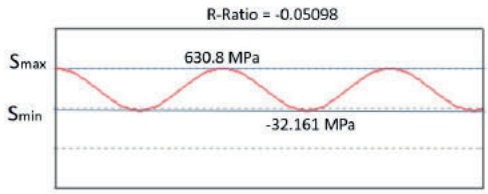


Figure 5: Determination of R-Ratio value for 0.01 mm shrinkage allowance

Analysis type was selected as "Stress Life". Because the strain life method is used in low cycles, while the stress life method is used in high cycles.

The Soderberg criterion was used as the average stress theory. Because it is the average stress theory that accepts the yield strength as the limit value. The Soderberg criterion curve is shown in Figure 6 [12]. The most critical part in layered cylinders is the inner cylinder. When deformation occurs in the inner cylinder, it loses its function.

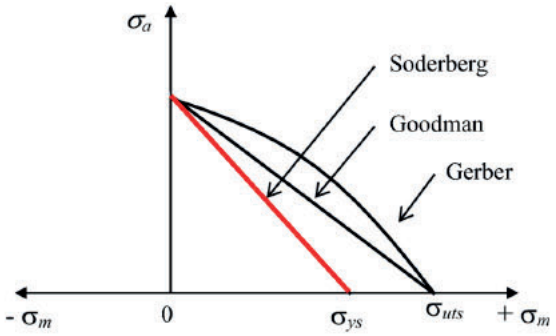


Figure 6: Gerber, Goodman and Soderberg lines [12]

Fatigue analysis were carried out for 11 shrinkage allowance values from 0.01 mm to 0.11 mm and the effect of shrink-fit on fatigue was revealed.

III. Results and Discussion

The life curve against the shrinkage allowance obtained as a result of the analysis is given in Figure 7.

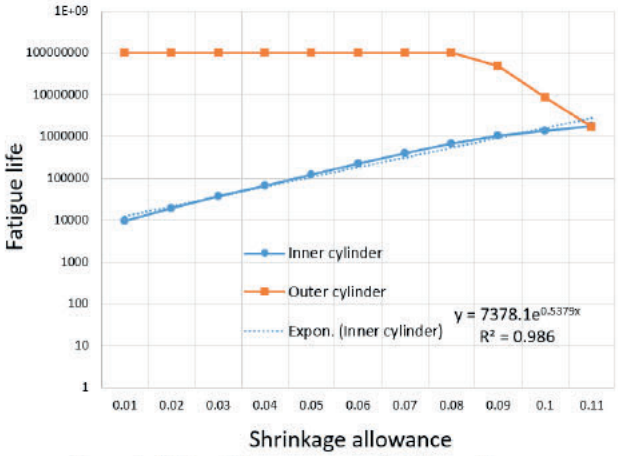


Figure 7: Fatigue life curve versus shrinkage allowance

As the shrinkage allowance increased, the pre-stress on the inner cylinder increased. The equivalent stress on the inner cylinder decreased as the pre-stress increased. As a result, the life of the inner cylinder has increased. Values above the 0.11 mm shrinkage allowance have not been studied. The reason for this is that R-Ratio values are above -1. When the shrinkage value increased by 1 percent in the inner cylinder, the life of the cylinder increased almost linearly between 1.6 and 2 times. However, the life of the outer cylinder suddenly decreased because the shrink fitted inner cylinder applied high force on the outer cylinder.

Life curves of the inner and outer cylinder intersected at approximately 0.11 shrinkage allowance. When low shrinkage allowance is used in systems, the outer cylinder can be safe, but the inner cylinder that sees and transmits the pressure will quickly fatigue. Instead, by giving applicable large tolerances to cylinder geometries, both cylinders can be operated for a very long life.

The fatigue life obtained for a shrinkage allowance of 0.01 mm is given in Figure 8a. In Figure 8b, the fatigue life obtained for 0.11 mm shrinkage allowance is given. For 0.01 mm, the life of the inner cylinder is 9530 cycles, while the life of the outer cylinder is 10^8 cycles. For 0.11 mm, the life of the inner cylinder is 1808100 cycles, while the life of the outer cylinder is 1685000 cycles. Life increased by 189.73 times in the inner cylinder and 59.35 times in the outer cylinder.

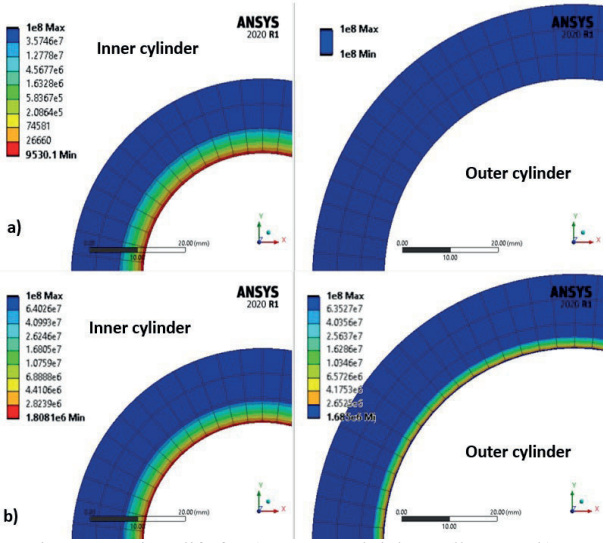


Figure 8: Fatigue life for a) 0.01 mm shrinkage allowance b) 0.11 mm shrinkage allowance

The fatigue life of a single layer cylinder of the same geometric size and a double layer cylinder with a shrinkage allowance of 0.01 mm was investigated. While the life of the single-layer cylinder is 5586 cycles (Figure 9), the life of the double layer cylinder with a shrinkage allowance of 0.01 mm was determined 9530 cycles (Figure 8a). Single layer cylinder has 1.7 times less working life than double layer cylinder (cylinder with the least life) with the lowest shrinkage allowance.

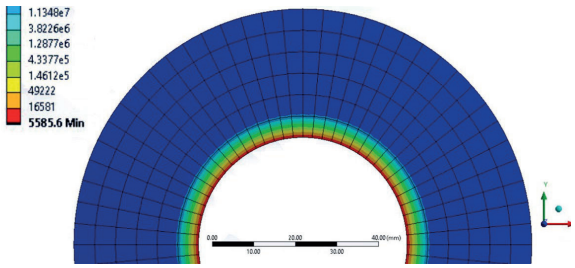


Figure 9: Fatigue life of single layer cylinder

S-N diagrams according to different stress ratios were shown in Figure 2. According to this diagram, as the stress ratio approaches -1, the fatigue life of the material and the life values increase. The analysis results given for the inner cylinder in Figure 10 also confirm the S-N curves. The life of the inner cylinder gradually increased as the stress ratio approached -1. The same change applies to the outer cylinder. As the stress rate increases, a decrease in outer cylinder life has been observed.

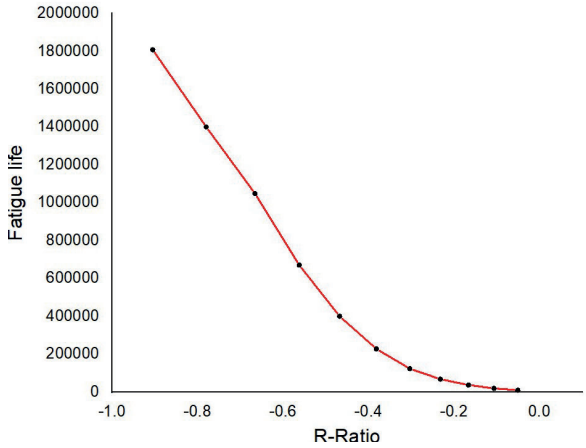


Figure 10: Fatigue life curve versus R-Ratio

IV. Conclusion

In this study, the fatigue life of compound cylinders depending on the shrinkage allowance and stress ratio is calculated. As a result, it has been determined that the shrinkage allowance positively affects the fatigue life of the compound cylinders. By choosing the optimal shrinkage allowance for the cylinder geometries, the fatigue life of the cylinders can be improved. The paper can contribute to future pressure about high-pressure fatigue tests.

Acknowledgment

In this study, material properties of AISI 4140 steel and Stress-Life Graphs based on different stress ratios were taken by Ces Selector Onatus Öngörü Teknolojileri.

References

[1] Ozturk, F Finite-element modelling of two-disc shrink fit assembly and an evaluation of material pairs of discs Proc Inst Mech Eng Part C J Mech Eng Sci 225, 263–273, March 2011

[2] Truman, C E & Booker, J D Analysis of a shrink-fit failure on a gear hub/shaft assembly Eng Fail Anal 14, 557–572, March 2007

[3] Patil, S A Fatigue analysis of compound cylinder subjected to repeated internal pressure Am Soc Mech Eng Press Vessel Pip Div PVP 2006, 1–5, July 2006

[4] Özel, A , Temiz, Ş , Aydin, M D & Şen, S Stress analysis of shrink-fitted joints for various fit forms via finite element method Mater Des 26, 281–289, September 2005

[5] Miraje A A , and Patil, S A , Infinite fatigue life of three layer shrink fitted compound cylinder under fluctuating internal pressure International Journal of Mechanical Engineering and Technology, Volume 3, Issue 1, 299-310, January 2012

[6] Özlü, T Mikro Modüllü Dişli-Mil Sıkı Geçme Bağlantılarında Mil Çapı ve Delik Toleransının Gerilme Dağılımına Etkisi *Eur. J. Sci. Technol.* 330–335, October 2019
[7] Chouchane, M , Fakhfakh, T , Daly, H Ben, Aifaoui, N & Chaari, F Design and modeling of mechanical systems - II: Proceedings of the sixth conference on design and modeling of mechanical systems, Hammamet, Tunisia 213–220, March 2015
[8] Paredes, M , Nefissi, N & Sartor, M Study of an interference fit fastener assembly by finite element modelling, analysis and experiment *Int. J. Interact. Des. Manuf.* **6**, 171–177, February 2012
[9] Marouani, H & Hassine, T Finite element investigations of the shrink-fit assembly with radial cyclic load *Lect. Notes Mech. Eng.* **789**, 213–220, 2015
[10] Gaul, L & Schmidt, A Finite element simulation and experiments on rotor damping assembled by disc shrink fits *Mech. Syst. Signal Process.* 127, 412–422, November 2019
[11] Ansys Granta Selector, Onatus Öngörü Teknolojileri
[12] González-velázquez, J L Mechanical Behavior and Fracture of Engineering Materials, 2019
[13] SolidWorks, Solidworks, Dassault Systems, Version 2020
[14] ANSYS 2020 Student Version

An Approaching for the Production of Special Profiles/Rifling's with Plastic Deformation

O.ACAR[1], M. KUNTOĞLU[1], H. SAĞLAM[1]

[1] Selcuk University, Konya/Turkey, osmanacar@selcuk.edu.tr, mkuntoglu@selcuk.edu.tr, hsaglam@selcuk.edu.tr

***Abstract* -In firearm barrels, several manufacturing methods have been used for increasing precision of shot with generating rifling marks to internal face of barrel and providing longer barrel life. Flow forming (tube spinning) is a preferred approach for producing cylindrical parts which enables improved surface finish and surface hardening because of cold forming (flow forming is performed at 25°C). In this study, a design providing the manufacturing of both rifling marks and barrel with plastic deformation using a mandrel possessing male form of rifling for the purpose of generate internal form of barrel is presented. On the contrary to traditional spinning wall thickness is reduced deliberately with ironing, this process is occurred with increased deformation despite less consumed force. The design aims to determine metal flowing rate and ensures the dimensional accuracy with avoiding expansion in barrel hole after shaping and thanks to this disconnect the mandrel with hole easier. Besides, compared to rifling forms shaped with machining operations, it is aimed that to obtain both faster and a surface possessing high strength and wear resistance in consequence of plastic deformation. Grinding with rollers provides extraction of rifling form to internal surface of barrel which enhances the dimensional precision. With this method, high strength and robust structure is obtained which provides longer barrel life and going out of bullet faster from barrel which increases the accuracy of shot is purposed**

***Keywords* – Rifling production, Plastic Deformation, Flow Forming, Tube Spinning.**

I. INTRODUCTION

RIFLING is performed by several manufacturing methods for increasing precision of shot with generating marks to internal face of barrel and providing longer barrel life. Flow forming (tube spinning) is a preferred approach for producing cylindrical parts which enables improved surface finish and surface hardening because of cold forming (flow forming is performed at 25°C). In this study, a design providing the manufacturing of both rifling marks and barrel with plastic deformation using a mandrel possessing male form of rifling for the purpose of generate internal form of barrel is presented. On the contrary to traditional spinning wall thickness is reduced deliberately with ironing, this process is occurred with increased deformation despite less consumed force. The design aims to determine metal flowing rate and ensures the dimensional accuracy with avoiding expansion in barrel hole after shaping and thanks to this disconnect the mandrel with hole easier. Besides, compared to rifling forms shaped with machining operations, it is aimed that to obtain both faster and a surface possessing high strength and wear resistance in consequence of plastic deformation. Grinding with rollers provides extraction of rifling form to internal surface of barrel which enhances the dimensional precision. With this method, high strength and robust structure is obtained which provides longer barrel life and going out of bullet faster from barrel which increases the accuracy of shot is purposed.

It has been increasingly significant for manufacturers to improve the technology using for firearms for enhanced quality, robust design and less wastage material and time. Longer lifetime and accurate targeting identify the robustness and reliability of a gun. Meeting these expectations highly depend on the shaping methodology for both internal and external surfaces. Advantages and disadvantages of an approach push the researchers for optimum results and make the production faster than the previous studies. The straightness of the barrel and rifling increases the quality of shot which directly refers to the selected manufacturing method. Meanwhile, surface texture and mechanical properties have determinant effect on barrel life. The producing method become important by means of the former said two items which will be the main aim of this research.

The manufacturing of barrels started with the raw bars then using drill bits, a hole is made in the material. To produce a perfectly shaped cylinder, reamer tools are took in consideration and additionally honing or lapping may be implemented for final desired dimension. In the past, hook rifling, broach rifling, hammer forging, button rifling, electrochemical rifling and casting were preferred for execute the rifling in barrel (Monturo, 2019:143) (Figure 1). A generally known way of producing rifling is broaching which has the main disadvantage of chipping causes weak finish surfaces (Figure 1D). A similar and older way with cutting for constitution of rifling is hook rifling which has single cutter to remove material (Figure 1B). This approach is preferred rarely in modern day defense technologies because of the time-consuming structure of single tool can cut one groove in a single pass. Hammer forging is a cold working process benefit from the hammers surrounding the raw material and a mandrel has a reversed geometry of rifling marks was used to give the form to the barrel (Figure 1A). High cost equipment and hard to determine the measurements of grooves and lands lead to decrease of designation of this method. Button rifling is performed with tools named button which has reversed lands

and grooves on mandrel (Figure 1C). Button rifling is a cold working process and also a cutting mechanism is performed lighter than broaching. Electrochemical rifling is a modern way to produce rifling with electrolysis which has a shaped cylinder as lands and grooves (Figure 1E).

In the past, authors presented two novel approaches for high quality shaped rifling and barrel with minimum material loss, having wear resistance and high strength surface, for enhanced lifespan of tool (Kuntoğlu and Sağlam, 2017:1, Kuntoğlu and Sağlam 2017:1). Calderon et. al. has performed a mechanical design and manufacture of a newly developed mandrel (Calderon et. al. 2012:76). Matsuzaki et. al. proposed an analytical model and experimental investigation to stabilize the chatter vibrations in deep hole drilling process during generating rifling marks (Matsuzaki et. al. 2015:194). On the other study, 3D modelling of tube rifling operation was performed, and optimum conditions have been found (Bella et. al. 2017:1382). This study contains a newer design approach for producing barrel body and rifling marks using flow forming operation which has a mandrel shaped male form of rifling. Easily imparting of mandrel to internal surface of barrel provides to manufacture of barrel in desirable tolerance ranges.

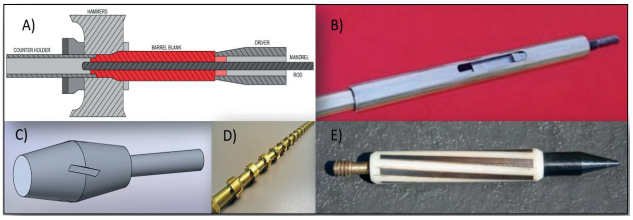


Figure 1. Several rifling producing methods

II. MANUFACTURING RIFLINGS WITH FLOW FORMING METHOD

Flow forming or tube spinning is like shear forming and generally preferred to form the cylindrical parts in manufacturing operations. During operation, a rotating mandrel pioneered the workpiece with the aid of two or three rotating rollers and material is exposed to plastic deformation. Axially replaced material is getting thinner while the length is growing longer through the mandrel body (Wong et. al. 2003:1430). Getting inspired from flow forming, a novel design is presented to adopt flow forming for producing a rifled barrel. Forward or backward type of flow forming can be experienced for optimal solution according to manufacturer's option. A general approach, design and discussion is carried out in this study.

Before flow forming operation, a male contoured rifling should be generated on a cylindrical sample. The executed force composes a plastic deformation result in yielding of the material to external and internal surface of the mandrel. During the material flow, the grooves between workpiece and mandrel are filled to reproduce the lands of the rifling. A male form of rifling has been imparted to internal surface of mandrel (Figure 2). The workpiece and mandrel can be fixed to headstock to perform the operation on a spinning or a shear forming machine. With feed of rollers, metal flows to reverse direction and the preform of the workpiece transform the newly produced component. The rough forming can be performed with the replication of flow forming operation. After cold working, rollers can be delicately getting through the workpiece several times to realize the ironing operation providing the finish shaping. Unlike spinning, the material yields through the mandrel and rifling body, the thickness of the wall is decreased. A possible difficulty can appear after flow forming operation is dilatation because of cooling down. With regulating the material flow rate seamlessly, desired dimensions can be acquired and also, dilatation effect provides to pull out the mandrel from workpiece. By means of plastic deformation mechanism, the rifling form is transferred as a reversed copy and the obtained structure enables high wear resistance, longer bearing life and high strength. Revolving and axially moving rollers have grinding effect on material which easily copy the rifling profile on workpiece. The internal surface of barrel can be shaped with improved dimensional accuracy.

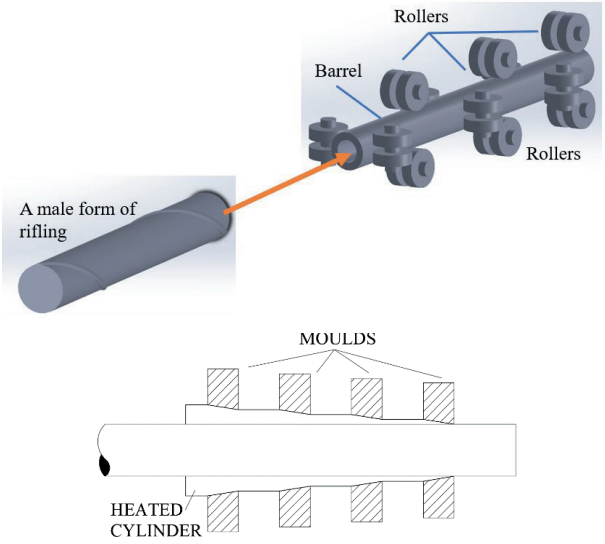


Figure 2. The Flow Forming Operation

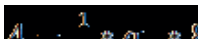
III. PLASTIC DEFORMATION AND DEFORMATION ENERGY

A material being exposed to stress, elastic and plastic deformation occurs according to the material properties i.e. ductility or fragility which depend on elasticity module, crystal structure, atomic bond and bond energy. The general structure of the shaping methodology in flow forming is based on plastic deformation. The area under the curve on stress-strain graphs identifies the deformation energy or toughness which can be defined as the capability of energy absorption of a material and ability for plastic deformation. Material

properties, elasticity modulus, heat treatment, alloying element, crystal structure have influenced on deformation energy. The calculation of deformation energy is derived from the stress-strain graphs which is the absorbed energy of a material for unit volume. The deformation energy can be calculated as following equations:

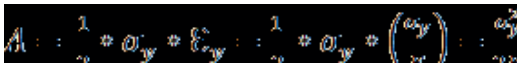
$$A = \int_0^{\varepsilon_y} \sigma * d\varepsilon \quad (1)$$

Where A is the area under stress and strain graph, Ɛy is unit deformation in yield, $\boldsymbol{\sigma}$ is stress in equation 1. Considering the elastic area is linear, then the equation can be calculated as;



(2)

Considering the $\boldsymbol{\sigma}$ = E * Ɛ then, the equation 2 is transformed to equation 3 where the obtained result is toughness or deformation energy according to SI system the unit of it is J/m3 or Pa.



(3)

IV. Conclusion

Production of barrels and manufacturing of rifling become an important task in today's developing defence industry. The requirement for high quality product demonstrates nations power in international community. For this purpose, a technological contribution is purposed in this paper for firearm manufacturing.

One of the purposes of the paper is to develop a new approach for manufacturing of rifling barrels with higher quality. By this way, desired wall thickness for barrel and rifling profile in seamless dimensional can be achieved.

The design can be implemented easily to any male form rifled body with flow forming approach. Although the preprocess take time in designing the mandrel according to selected material, machine tool and dimensional content, possible for fast and favorable mass production.

This approach revives the importance of manufacturing method which increase barrel life with high strength and wear resistance as a result of plastic deformation mechanism. Cold working eliminates the disadvantage of shaping with cutting and different from other cold working approaches less material loss and faster production can be obtained.

References

[1] Bella, P., Bucek, P., Ridzon, M., Mojzis, M., Parilak, L., (2017). "On numerical modelling of multi-rifled tube drawing ", International Conference on the Technology of Plasticity, 17-22 September, Cambridge, United Kingdom.

[2] Calderón, G. U., Urriolagoitia-Sosa, Verduzco-Cedeño V. F., Hernández-Gómez, L.H., Romero-Ángeles, B., Torres-Torres, C., Torres-San Miguel C.R., Wong-Angel, W.D. (2012). Mechanical design and manufacture by powder technology of a new mandrel tool for the fabrication of assault rifle barrels. Powder Metallurgy, 224, 76-85.

[3] Kuntoğlu, M., Sağlam, H. (2017). "The Production of Rifling Marks with Metal Spinning", 25th International Conference on Composites of Nano Engineering, 16-22 July, Rome/Italy.

[4] Kuntoğlu, M., Sağlam, H. (2017). "An Alternative Method for Forming of Rifling Groove on Firearms: Ball Rolling", 2. International Defense Industry Symposium, 06-08 April, Kırıkkale/Turkey.

[5] Matsuzaki, K., Ryu, T., Sueoka A., Tsukamoto, K., (2015). Theoretical and experimental study on rifling mark generating phenomena in BTA deep hole Drilling process (generating mechanism and countermeasure). International Journal of Machine Tools & Manufacture, 88, 194-205.

[6] Monturo, C. (2019), Forensic Firearm Examination, *Barrel Manufacturing and Rifling*, *Vol. 6* (pp. 143-156). Academic Press.

[7] Wong, C. C., Dean, T. A., Lin, J., (2003). A Review of Spinning, Shear Forming and Flow Forming Processes, International Journal of Machine Tools & Manufacture, 43, 1419-1435.

Investigation of Temperature Distribution in a Cooking Chamber of an Industrial Oven

A.H. ALTUN[1], S. U. ATMACA[2], H. A. EROL[3] and S. DOGAN[2]

[1]Selcuk University, Konya/Turkey, ahaltun@selcuk.edu.tr
[2]Konya Technical University, Konya/Turkey, suatmaca@ktun.edu.tr
[3]FIMAK Bakery and Gastronomy Machines, Konya/Turkey, huseyinerol@fimakmakina.com.tr
[2]Konya Technical University, Konya/Turkey, sercandogan@ktun.edu.tr

***Abstract* -It is demanded that and the final product should be delicious and the quality should be at a proper consistency when cooked in an industrial oven. Thus the temperature distribution needed to be diffuse homogeneously in the cooking chamber. According to these a food cooking oven selected, used for research and development in a business corporation, coded ESTONE 120*80. The temperature distributions are investigated experimentally by using a test setup. An experimental set-up is used for investigating the temperature distribution. The thermocouples are placed at 16 points for taking data during the cooking operation. It is also seen that the temperature distribution is almost homogeneous in the cooking chamber. By this, it is concluded that the temperature distribution can be investigated truly by this kind of an experiential set-up.**

***Keywords*– Cooking, heat transfer, industrial oven, temperature distribution, thermocouple.**

I. Introduction

There is a large amount of works concerning about the industrial ovens in the literature. Some of them can be given as; CFD modeling of an electric oven cavity to simulate cook and broil cycles; stably stratified flow being a characteristic of the latter has been discussed by Mistry et al. (2006). At the result of his paper a comparative analyses of cook and broil cycle thermal fields inside the oven is given.

A research concerning about reducing the energy consumption at ancestry electric oven, made by Hancilar (2007). The purpose of this work was raising the energy class of an oven from level C to level A. It is shown in the manuscript that the energy consumption reduced by 30% for the electric ancestry ovens.

A study made for determining the effect of some dominant parameters at ventilating the domestic type ancestry ovens to energy performance by Pehlivan (2010). This work constitutes a step for future researches for interpretation of thermal performance of the ovens according to different kind of criteria's. At his paper (Bödöşov, 2011) developed a mathematical model for heat and mass transfer for round cake cooking. In his research he simulates the cooking of products which is the most important process in this area. A thermal analysis of electric oven is performed by Ploteau et al. (2012). Not only an experimental set up is used but also finite element method is used for determining numerical model for this oven. Comparisons were made and similarities were shown at the results for numerical and experimental work. A new model is constituted by Ramirez-Laboreo et al. (2016) for evaluation of system performance of electric ovens. Simulations were made for different kind of cooking methods and distributions of energy and time results are given. Kişin et al. (2019) aimed to take maximum efficiency with minimum electricity and tried to determine a homogeneous temperature distribution at multi tray industrial ovens. At their papers a detailed knowledge about design of inlet and suction rims, blowers and power of electric resistant are given.

In this paper an experimental investigation is made for determining the temperature distribution of cooking chamber, on the manufacture of the a industrial oven, used for research and development in a business corporation, named ESTONE 120*80.

II. Experimental Setup

Temperature measurements are taken for improving the energy conception of an oven named as ESTONE 120*80, used for cooking of food product. The dimensions of the oven are 120x80 cm and still actively used for cooking. A measurements set-up is prepared for taking the values of temperatures from 16 nodal points inside the cooking chamber. The thermocouples are placed in a newly designed measurement system as a matrix form. Calibrations of the thermocouples were done for accurate measurements. The experimental measurement system shown in Fig. 1 placed in the cooking chamber.

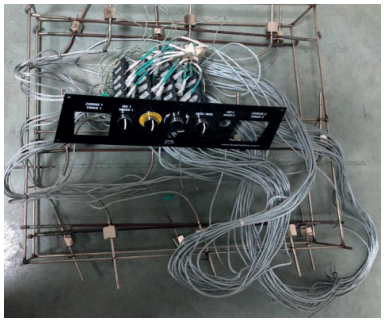


Figure 1: Test setup for temperature measurements.

The locations of the thermocouples were chosen as the critical points. The transient temperature distribution is

recorded by a data-logger system. After obtaining the result from the data-logger system, the results were processed.

The upper and lower side resistances were used for heating the oven chamber. Before the experiments the cover of the chamber kept open for ensuring the equality of oven chamber temperature and the ambient temperature. The inside temperature of the chamber is adjusted to the value of 230°C by a control unit after closing the door of chamber. This temperature value is chosen as the proper value while cooking the products. The oven heated to 230°C approximately 1.5 hour. During this process the transient temperature variations observed by a computer monitor. The temperature values are taken by thermocouples, placed parallel to the front cover of the chamber. Thermocouples are placed on the two side walls of the oven as 6 and 4 in the middle as shown in Fig. 2.

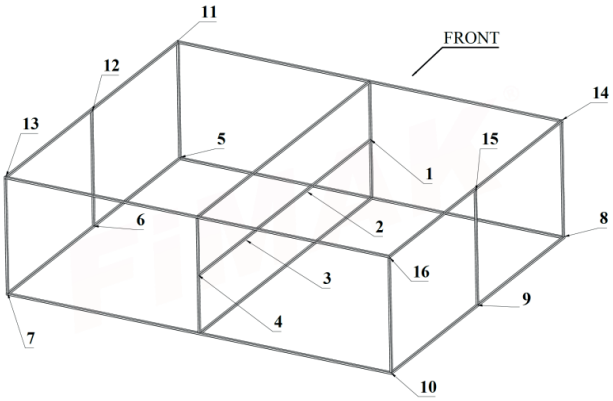


Figure 2: Schematic view of the test setup inside the oven.

III. Results and Discussion

The inside volume temperature of the oven chamber of ESTONE 120*80 is increased from ambient temperature to the cooking temperature transiently. A measurement system which is specially prepared for this oven is used which is arranged with 16 thermocouples as a matrix form. The temperature values at critical locations inside the chamber recorded depending upon time. The temperature values and time variation of different rows are shown at Figs. 3, 4 and 5.

The first row of the temperature measurement system has 4 thermocouples and placed in the middle of the oven. The 6 thermocouples which are located at second row and the third row with 6 thermocouples placed to the side walls of the chamber. The temperature values are recorded from whole nodal thermocouple points. The temperature variation inside the oven chamber simulated as cooking of a product. The inside temperature of the chamber is set to a value of 230°C by a control panel. The oven cover kept open for 60 seconds and closed again as a simulation. A sharp temperature decrease observed due to cover push-pull operation. Than the electric resistant cut-in and it is waited to get the desired temperature (230°C). This scenario can be seen from the temperature –time figure. A sharp temperature decrease occurs approximately at 90th minute.

Vapor addition is used in cooking operation for a bright product shell and soft structure at a closed chamber for a certain time. After push-pull operation of the oven cover it is waited to reach the temperature vale of 230°C. At the experiments for simulate as a real time cooking, at the 106th minute vapor is added to the chamber for 5 seconds period. The temperature decrease due to vapor addition can be seen from the figures. After the electric resistant cut-in again and the oven chamber is left to get desired temperature. The temperature measurements recorded during 120 minutes period and the experiment is ended.

The temperature values according to time for four thermocouples are shown at Figure 3. As can be seen from the Fig. 3 the maximum temperature difference among the four thermocouples is 10°C. The desired temperature value at the oven chamber is attained by blue colored graphic firstly at about 60 minutes. Minimum values are corresponds to the red colored thermocouple. The desired temperature values are obtained by the whole thermocouples for each experiment. And the values are so closer to each other at the end of the 120 minutes period.

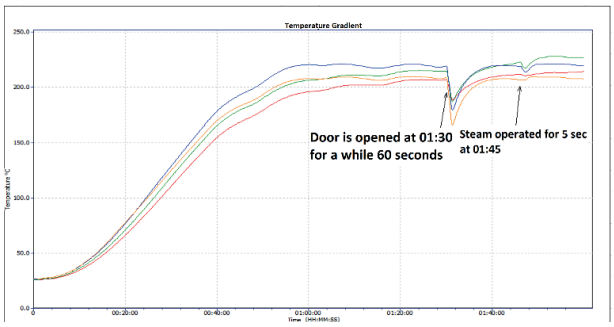


Figure 3: Change of temperature and time for four thermocouples in the middle of the oven.

The six different temperature values with time are given at Fig. 4 for the left side wall of the oven. The nearest values to desired temperatures are obtained by red and orange colored thermocouples. As can be seen from the figure, the final temperature values are so closer to desired temperature. And also the green and purple colored thermocouple values are closer to the desired temperature. The maximum temperature difference occurred at black and blue colored thermocouples as approximately 5°C.

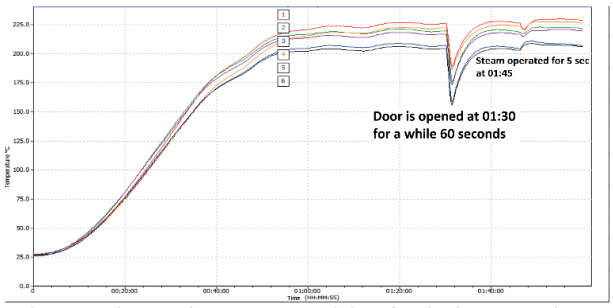


Figure 4: Change of temperature and time for six thermocouples at left side wall of the oven.

It is evaluated as it is normal to have a difference like these values due to higher desires temperature value and also to have a large chamber volume. A further job plan is scheduled to reduce the temperature value differences among the

thermocouples and it is demanded a more homogeneous temperature distribution in the oven chamber.

At the Fig. 5 the temperature values obtained from 6 different thermocouples placed at the right side wall of the chamber (3rd row) is shown. According to these values it can be seen that the needed duration to obtain the desired temperature value is approximately 60 minutes. And also it can be seen from Fig. 5 that all the thermocouples get the same temperature values. It can be concluded that this local region in the oven chamber is more homogeneous according to the temperature values than to the other parts.

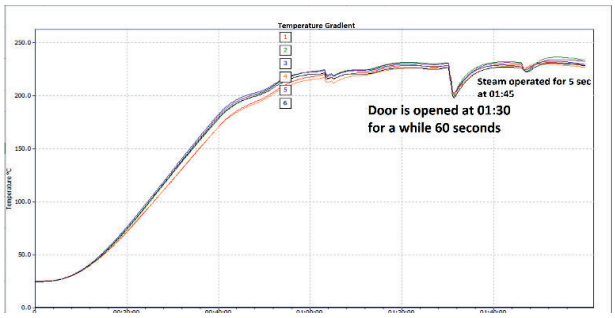


Figure 5: Change of temperature and time for six thermocouples at right side wall of the oven.

IV. Conclusion

A test setup was prepared to determine the temperature distribution in an industrial bakery oven. For temperature measurements, 16 thermocouples are placed at critical points in the oven. Temperature measurements were made taking into account different scenarios such as setting the ambient to the desired temperature in order to cook a food product in the oven, placing the food product in the oven and giving steam. According to the test results; the changes in temperature distribution in the oven during a cooking period and whether the resistances heat the ambient homogeneously are interpreted. From the temperature values taken, it was evaluated in which part of the oven the temperature difference was greater and which region reached the desired temperature. Some of the results obtained from experiments are given below.

• According to the temperature values depending on the time; it has been determined that the interior of the oven heats up to the desired temperature with a difference of approximately 3-4°C.

• According to the temperature values taken from different critical regions the interior of the oven; it has been observed that there is a difference of around 10°C especially in the regions in the middle of the oven.

• According to the temperature data taken from the left side wall of the oven with 6 thermocouples, it was determined that there is a more homogeneous temperature distribution compared to the middle region.

• It was determined that 6 different temperature values taken from the right side wall part of the oven were the same with each other and it was observed that the temperature was evenly distributed in the right side wall region.

• Consequently, the temperature distribution in the interior of an industrial oven (ESTONE 120*80) was scanned for different critical points and areas that require improvement have been determined.

• The obtained results can be used for validation in numerical and analytical studies.

Acknowledgment

This study was carried out within the scope of the R&D Center studies in FIMAK Bakery and Gastronomy Machines.

References

[1] H Mistry, Ganapathi-subbu, S Dey, P Bishnoi and J L Castillo, "Modeling of Transient Natural Convection Heat Transfer in Electric Ovens," Applied Thermal Engineering vol 26, pp 2448–2456, February, 2006

[2] N Hancılar, Experimental And Theoretical Investigation Of Energy Consumption For Electrical Domestic Built In Oven, M Sc, Graduate School of Natural and Applied Sciences Istanbul Technical University, 2007

[3] E Pehlivan, Investigation Of Ventilation Parametres On Thermal Performance On Build-In Domestic Ovens, M Sc, Graduate School of Natural and Applied Sciences Istanbul Technical University, 2010

[4] A Bödöşov, Mathematıcal Modelıng For Heat And Mass Transfer Durıng Rıng (Round) Cake Cooking In A Convectıve Oven, M Sc, Graduate School of Natural and Applied Sciences, Kyrgyz-Turkish Manas University, 2011

[5] J P Ploteau, V Nicolas and P Glouannec, Numerical And Experimental Characterization Of A Batch Bread Cooking Oven, Applied Thermal Engineering, Vol 48, pp 289-295, 2012

[6] E Ramirez-Laboreo, C Sagues and S Llorente, Dynamic Heat And Mass Transfer Model Of An Electric Oven For Energy Analysis, Applied Thermal Engineering, vol 93, pp 683–691, 2016

[7] B Kişin, M Burhan, O Dede, K İleri and C Çelik, A Alper ÖZALP, Fırın İçi Momentum Ve Isı Transferi Mekanizmalarının Numerik İncelenmesi, 14 ULUSAL TESİSAT MÜHENDİSLİĞİ KONGRESİ, 17-20 NİSAN 2019/İZMİR, pp 1127-1147

A Research on the Door Lock system of Armored Military Vehicles Using Hydraulic And Pneumatic Components to work coupled with High Strength Steel

M. İŞKUR[1], F.BEDİR[2] and A.ÇOKAL[1]

[1] Şener Mechanic Systems Industry and Trade Inc. İstanbul/Turkey mesutiskur@senermekanik.com,
[2]Gebze Technical University, Kocaeli/Turkey, fevzibedir@gtu.edu.tr
[1] ŞENER Mechanic Systems Industry and Trade Inc. İstanbul/Turkey ahmetcokal@senermekanik.com

Abstract - In this study, a research has been carried out on a locking system that minimizes external vibrations and impacts in 4x4, 6x6, 8x8 armored vehicles in the automotive and defense industry. The parts used in the system are supported by high-strength steels and a new lock mechanism design is envisaged where the hydraulic piston and pneumatic piston are used at the same time. Static and dynamic parameters that may occur under external forces that the vehicles are exposed to will be foreseen. Static structural analysis will be performed.

Keywords **- Lock mechanism, Military vehicle, Hydraulic and pneumatic components, static analysis, design parameters.**

I. INTRODUCTION

As in the past, today armored vehicles are used in almost all military vehicle systems in the military field. With advancing technology and increasing needs, military armored vehicles have started to be produced in keeping with today's conditions and application areas by developing and improving many different features. Military wheeled armored vehicles are produced in a way that is fit for difficult terrain conditions and are produced by taking into consideration all kinds of climatic conditions with their high mobility increasing their efficiency and their speed in long-distance road conditions. Moreover, these vehicles, custom-designed to meet the needs of users, may come with many various features and attributes, such as 4x4, 6x6 and 8x8, and there are many different examples thereof. Being products of Turkish Defense Industry, vehicles such as Amazon 4x4, Arma 6x6 and Arma 8x8, Kirpi 4x4, Kirpi 6x6, Akrep II 4x4, Hızır 4x4, Ejder Yalçın 4x4, Ejder Kunter Special Purpose Platform 4x4- 6x6, Cobra I and Cobra II 4x4, Pars 4x4, Pars 6x6 İzci, Pars III 6x6, Pars III 8x8, Pusat 4x4, Ural 4x4, Vuran, Yörük 4x4, Arma 6x6, Armored Ambulance 4x4 can be cited as examples of wheeled armored vehicles.

The door lock system in all of our existing military land vehicles is designed with bending parts connected to a welded handle mechanism. In such design, the axial misalignment occurring upon sheet plate bending as well as the distortions stemming from welding hinder comfortable and secure use of the door system. Door lock system or with another definition latching system and other components are disclosed that are particularly well suited for use with the heavy doors of armored military vehicles. Some component embodiments are usable in normal and emergency modes to provide access, entry, egress and rescue through vehicle door openings [1]. Automotive liftgate latches have been subject to regulation for minimum strength and inertial resistance requirements since the late 1990's in the US and globally since the early 2000's, possibly due to liftgate ejections stemming from the first-generation Chrysler minivans which employed latches that were not originally designed with this hazard in mind. Side door latches have been regulated since the 1960's, and the regulation of liftgate, or back door latches, have been based largely on side door requirements, with the exception of the orthogonal test requirement that is liftgate specific [2]. Salmon and Eikey showed in their study to determine the effects of friction on the systems dynamic response by some physical tests and analysis of a typical automobile latch and outside handle release mechanism. The results of the test show that friction forces internal to the latch significantly affect the quasistatic and dynamic behavior of the latching system [3]. Nelsen and Seo performed calculations to verify the compliance of automotive door latch systems to minimum legal requirements as well as to perform additional due diligence calculations necessary to comprehend special cases such as roll over crashes and locally high inertial loadings. [4]. Jain and Adhikari presented in this study, a methodology adopted to address the issue of latch failure evaluation and the critical observations made during the course of study. The verification of the developed methodology is done by correlation of CAE results with physical test results [5]. Karla et all. investigated side door latch performance in full vehicle and door system component testing as opposed to the component-only fixture testing outlined in FMVSS 206. Both full vehicle and door system testing showed that the tested side door latches managed forces that were substantially higher than any loads required by FMVSS 206 when subjected to purely vertical loading [6]. Gross studied on the factors to cover the

most significant causes relating to accidental door opening such as various types of loading of significance, description of five known causes of accidental door opening, supporting evidence from full-scale collision research [7]. The governments of many countries have established regulations that address the issue of vehicle door safety during crash events. Depending on the regulation or specification, analytical tools may be acceptable for verifying the crashworthiness of a latching system. In those instances, where actual test crashes are required for verification, analytical methods can still be used to help predict the outcome of a crash test [8].

The aim of this study is to offer a lock system for military land vehicles. It is thus to ward off the jamming of the door, provide the necessary security and defense conditions in military land vehicles in cases where the vehicle's armor withstands the impact during a possible mine attack in military operations.

II. EXPERIMENTAL PROCEDURE

In this study, it can be shown our suggested design of the door mine up/down lock of land military armored vehicles in Fig.1. The lock mechanism was analyzed by using ANSYS Workbench 2020 R2.

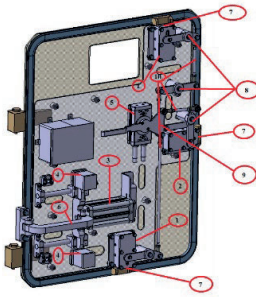


Figure 1: Our suggested 3D solid design of the door mine up/down lock of land military armored vehicles

It is given some detail of critical parts respectively in Fig. 1; Mine lock system (1) and the door lock system (2) were supported by the pneumatic actuator (5). In order to open the door after a possible explosion, the tension arm (6) carrying the load of the door is supported by hydraulic (4) and pneumatic piston (3).

It is accepted input data for the ANSYS model get from STANAG 4569 standard [9]. Total force required to open the lock system is 25kN, vibration resistance is max. 200 Hz and a load change is max. 25kN, the cycle number of opening-closing is 110000. It has been strengthened with 2 mine lock systems and 1 door lock system in the design. The door is connected to the vehicle body with two hinges from the outside. The door weight is 400 kg in armored form. The door is supported by a hydraulic-pneumatic system from the inside. The material descriptions used in the design are given in below Table 1. The surface pressure with the combination of carriage-spring-piston pusher enables the mechanism to engage and get locked to the fixed part of the door by making 50 mm x 20 mm pressure with its wide rectangular geometry.

Table 1 Materials used in the design

Item	Name	Materials
7	Lock Bolt	AISI 4340
8	Distributing Mechanism	34CrNiMo6
9	Handle system	17-4PH(630)
10	Drive Shaft	AISI 304

In this configuration of the prototype manufactured vehicle, the mechanism lock bolt is designed and improved from 4340 material. It is important for military security to ensure that the hub shaft at the center of the cover functions simultaneously with the locking mechanism inside. It will be ensured that the door can be easily opened by the military personnel inside the vehicle in the event of a possible strike/blow and that the cover retainer springs prevent the cover from falling over the personnel at the time of exit, thereby providing them a safe exit from the vehicle (For this, in the lock system, the bolt will be placed inside a special roller slot, and whereas the main handle system will ensure that the door opens with a roller and double piston instruction, which will constitute one of the most important security parts during a mine explosion.

III. RESULT AND DISCUSSION

It is shown in Fig.2 (a), for the known application, door opening-closing systems and mine lock systems are used in the form of a mechanical lock movement and as welded. In case of a possible mine attack, even when armors of the military land vehicles withstand the impact during the operations, an environment of security and defense cannot be created due to jamming of doors and inability to open them. What's more, since the vehicle is armored and heavy, the door hinges are far from performing the desired function and it is difficult to open and close the door of the vehicle due to the loads generated during a possible attack or while the vehicle is on the move, thereby putting the security of our military personnel at risk.

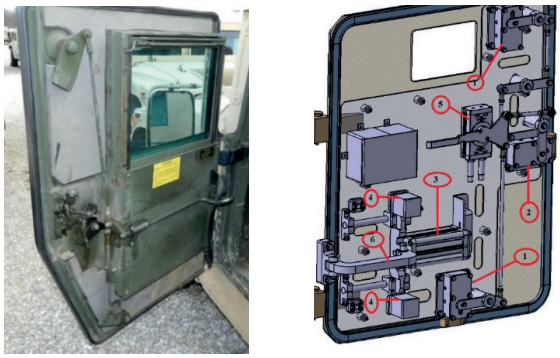


Figure 2: Military armored vehicle door systems: (a) A known application (b) Our design

As a result, the need for eliminating the drawbacks and deficiencies as well as the disadvantages of the constructions

and applications existing in common design and used today necessitates making a development in the technical field. Our present design (Fig.2-b) relates to a new military lock system developed so as to overcome the above-mentioned disadvantages and introduce new advantages to the relevant technical field as follows. It is to ward off the jamming of the door and thus to provide the necessary security and defense conditions in military land vehicles in cases where the vehicle's armor withstands the impact during a possible mine attack in military operations, to provide a door opening and closing system that operates via hydraulic and pneumatic movements, thereby protecting our military personnel from external threats in a possible attack and enabling an easy opening when they want to get out of the vehicle, to manufacture the materials to be used in the lock system from high-strength metal alloys, such as high-strength steels, tempered steels (AISI 4340 Tempering+ Induction Process), carburized steels (40CrNiMo6 to be carburized), and 17-4PH (SAE 630 Stainless Steel), to ensure a much more comfortable and easy locking and opening of the door by triggering via at least 2 piston systems at each door, and to ensure that the surface pressure with the combination of carriage-spring-piston pusher enables the mechanism to engage and get locked to the fixed part of the door by making 50 mm x 20 mm pressure with its wide rectangular geometry. Thus the door locking process is reinforced.

In the case of the suggested conditions, it is performed some analysis with ANSYS and obtained some results given in Figure 3. The distribution of force overcome to open door locks mostly affects the part in Figure 4. However, the analysis results showed that the maximum stress distribution was seen in the elastic region. It is predicted with the design that the maximum force required for the safe opening of the door can be increased up to three times.

Fig.3 shows the ANSYS statical structural analysis results of the lock system. The maximum stress occurred in distributed shell element, item 10 in Fig.1. It was resealable result for us because of transmitted 25kN applied on the door handle, item 9 in Fig.1.

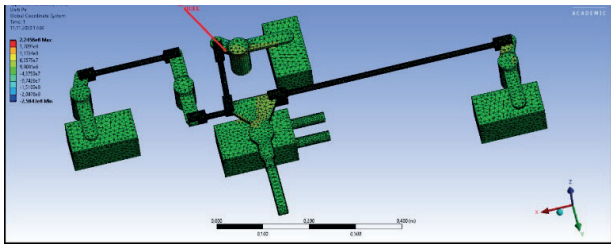


Figure 3: ANSYS statical structural analysis results of the lock system

Another important step of this study is that fatigue behavior of the lock system. It is understood from ANSYS analysis, the most critical part is item 10, made of the AISI 4340.

It well understands from literatures for the fatigue behavior of the AISI 4340 materials under nearly same conditions was done by Marcelino et all. [10]. The SN curves for axial fatigue tests result in their study shown in Fig.4.

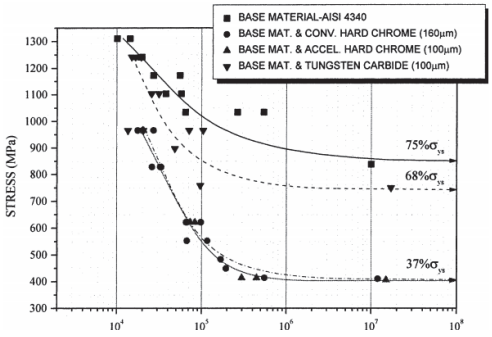


Figure 4: Fatigue behavior of the AISI 4340 materials done by Marcelino et all. [10]

The SN curves for the axial fatigue tests for the base metal AISI 4340, for the 110000 cycles number, fatigue life is shown in transition zone. It understands from this promising result for our lock-system must be revised and develop.

IV. Conclusion

Our present design relates to a new military lock system developed so as to overcome some disadvantages and introduce new advantages. In the case of loading conditions, it is run with the ANSYS to obtained critical stress value. The analysis results showed that the maximum stress distribution was seen in the elastic region. Obtained results from literatures for the fatigue behavior of the AISI 4340 materials shown that our offered lock-system must be revised and developed.

References

[1] United States Patent, Patent No : US 8,844,982 B2, Date of Patent: *Sep 30, 2014

[2] J Nelsen, H Wu, Crash-induced Loads in Liftgate Latching Systems, SAE Technical Papers, Published April 03, 2018 by SAE International in United States 2018-April

[3] J Salmon, C Eikey, The Effects of Internal Friction on Automotive Latch and Release System Behavior, Volume 2019-January, Issue January, 18 January 20192019 SAE Automotive Technical Papers, WONLYAUTO 2019

[4] J Nelsen, C S Seo, An Improved Methodology for Calculation of the Inertial Resistance of Automotive Latching Systems, Published April 01, 2014 by SAE International in United States

[5] TJain, T Adhikari, Latch Failure Prediction for Side Door Intrusion Test in CAE and Its Correlation with Physical Testing, Published January 14, 2015 by SAE International in United States

[6] K J Petroskey, M,E Klima, M P E M Paddock, Evaluation of Door Latch Response to Vertical Loading Conditions, Published April 20, 2009 by SAE International in United States

[7] A G Gross, Accidental motorist ejection and door latching systems, Automotive Engineering Congress; Detroit, MI; United States; 13 January 1964 through 17 January 1964; Code 90540)

[8] K P Jankowski, E Kamal, Vehicle door latch safety measures based on system Dynamics, 20042004 SAE World Congress; Detroit, MI; United States; 8 March 2004 through 11 March 2004; Code 90264)

[9] NATO Standard Agreement (Stanag 4569 standard) Draft, January 2004

[10] P Marcelino, R C Souza, W L Pigatin, H Voorwald, Effects of tungsten carbide thermal spray coating by HPHVOF and hard chromium electroplating on AISI 4340 high strength steel, Surface and Coatings Technology 138 2001 113-124]

Effects of Camelina Biodiesel Blends with Eurodiesel on Vehicle Performance

M. Mahmut YILDIZ[1] and A. Engin ÖZÇELİK[2]

[1] Selcuk University, Institute of Science, Konya/Turkey, mahmut.yildiz.4233@gmail.com
[2]Selcuk University, Faculty of Technology, Konya/Turkey, eozcelik@selcuk.edu.tr

Abstract **- Most of the world's energy needs are met from fossil sources. However, gases released by the gradual decline of fossil resources and the burning of fuels from these sources have negative aspects such as environmental pollution, acid rains or global warming. Due to such negative effects, studies are being carried out on alternative energy sources. Eurodiesel fuel used in diesel engines is the main source of energy for transportation and freight transportation. Biodiesel made from oils are important alternative fuels for diesel engines.**

In this study, camelina oil was used as a biodiesel raw material. Biodiesel production has been done by transesterification method. Biodiesel fuel obtained from camelina oil was mixed with diesel oil at the rate of 7% (B7), 20% (B20) and 50% (B50). Biodiesel-eurodiesel blends, four-stroke, four-cylinder, common-rail fuel system turbo-diesel engine, engine performance values were investigated on the chassis dynamometer. Looking at the results of the experiments, the highest engine moment among all fuels was obtained in B7 fuel at 80 km / h and engine power value at 160 km / h. The test results were compared with 100% eurodiesel fuel and interpreted, considering the performance values under different operating conditions.

Keywords **– Biodiesel, Camelina, eurodiesel, chassis dynamometer, vehicle performance**

I. Introduction

The unbalanced dissying of fossil fuels worldwide with the largest share of primary energy sources has also made some countries dependent on disadvantaged and advantageous countries[1]. This dependency is being further ed with the increase in energy demand. As a result, countries are now on the path of first looking for ways to evaluate their own self-energy sources, reducing dependence on foreign countries and diversification of energy resources to ensure supply security [2].

Biodiesel, canola, sunflower, corn, cotton, camelina etc. It is a product that is released as a result of the reaction of vegetable oils or animal oils obtained from oilseed plants such as oilseeds with methanol, a short chain alcohol, in the presence of a catalyst and used as fuel. Domestic cooking oils and animal based oils can also be used as biodiesel raw materials [3].

Camelina stands out as an important source of biofuel. The high oil content of camelina seeds (25–48%) [4] and (> 90% unsaturated fatty acids), together with the short life cycle of the plant, low agricultural input requirement, and its relative tolerance to various biotic and abiotic stresses [5]. Camelina is dried after harvesting and blending. Mechanical pressing in camelina is the main method of processing seeds, which can be carried out in different ways. Another method is to press the camelina seeds into roasted paste under mechanical pressure. In this method, the seeds are mixed with an equal amount of water to give the mixture a doughy appearance. Then the paste is roasted at 60-90 ° C and after the process, the mixture becomes sandy. The roasted paste is squeezed with a mechanical press and the oil is removed. Finally, the oil is filtered to clarify. In this study, it was not done in order to determine the performance of diesel and camelina biodiesel in a vehicle with a diesel fuel system with a vehicle dynamometer.

Table 1. Comparison of crude camelina oil and diesel fuel properties [6]

Specifications	Camelina crude oil	Diesel
Density 15 0C (kg/m3)	918	838
Kinematic viscosity 40 0C mm2/s	24	2.92
Flash point 0C	>220	102
Lower Calorific Value (MJ/kg)	38	42.3
Ash (% kütle)	0.0025	0.01
Sulfur (mg/kg)	13.85	9
Water content (mg/kg)	710	43.8
Acid value (mg KOH/g)	1.39	-
Iodine number g.I2/100 g)	151.5	-

II. Materials and Methods

Selcuk University Technology Faculty Mechanical Engineering Biofuel Laboratory was used to produce biodiesel from camelina oil. Selcuk University Faculty of Technology Automotive with camelina biodiesel obtained from the supplied camelina oil, a 2000 cc turbo-charged common-rail diesel fuel system, 2012 model Toyota brand Verso 2.0 type

vehicle was used. The technical specifications of the engine belonging to this vehicle are given in Table 2.

Figure 1: Experiment setup and vehicle

Table 2. Technical characteristics of the test vehicle

Brand	TOYOTA
Model	Verso 2.0
Drive shape	Front wheel drive
Year of manufacture	2012
Wheelbase	2780
Vehicle length / Width / height (mm)	4440/1790/1620
Maximum vehicle weight (kg)	1540
Vehicle engine characteristics	
Number of cylinders	4
Diameter x stroke (mm)	86 x 86
Engine volume (cm^3)	2000
Number of valves	16
Compression ratio	15:8
Fuel system	Diesel commonrail
Maximum power (HP-d / d)	124-3600
Maximum torque (Nm-d / d)	300-2400

The production of Camelina Biodiesel (Camelina Oil Methyl Ester) was carried out by the transesterification method in the Biofuel Laboratory in the Mechanical Engineering Department of Selcuk University Technology Faculty. Biodiesel fuel obtained from camelina oil was mixed with diesel oil at a rate of 7% (B7), 20% (B20) and 50% (B50). The viscosity, density, calorific value and flash point tests of B7, B20, B50, B100 and Eurodiesel fuel obtained as a result of the blends were carried out at the Biodiesel Laboratory of the Agricultural Machinery Department of the Selcuk University Faculty of Agriculture.

In the experiments, the HPT 6100 brand chassis dynamometer shown was used. The vehicle dynamometer used can perform road test, power test and speed test [7,8]. The technical information of the device of the Chassis Dynamometer to be used in the tests is given in Table 3 [9].

Table 3. Technical characteristics of the chassis dynamometer

Brand		HPT
Model		6100
Drum Diameter	m	1.2
Max Speed	km / h	200
Maximum axle load	kN	35
Capacity	kW	170

III. RESEARCH RESULTS AND DISCUSSION

The vehicle is 4. gear is fully throttled. The maximum engine speed is 6500 d/d and the speedometer is read at 160 km/h. When the message "press the clutch pedal" was read on the computer monitor, it was pressed into the clutch by the operator and the wheels were expected to stop on their own. The brakes should never be pressed during deceleration.

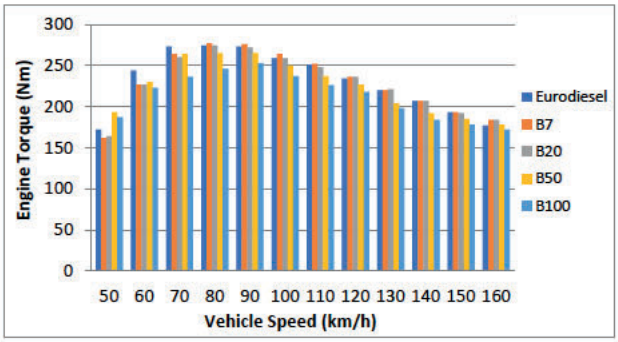


Figure 2. Engine moment change of fuels depending on vehicle Speed

Figure 2 shows the change of engine moments depending on vehicle speed for different fuels. The maximum engine moment values for The Euroiesel, B7, B20, B50 and B100 fuels were calculated as 274, 277,274, 265 and 253 Nm respectively. When looking at the results of the experiment, the highest moment value of all fuels was obtained at 80 km/h B7 fuel.

Eurodiesel fuel usage in terms of engine momentum was higher than B7, B20, B50 and B100 fuels. With the increasing percentage of biodies in the mixture, moment values decreased. Lower moment values are due to the lower thermal value of the biodieline than diesel fuel and higher viscosity. When the motor torque values are examined, it is seen that the work done by is similar to the motor torque values [9].

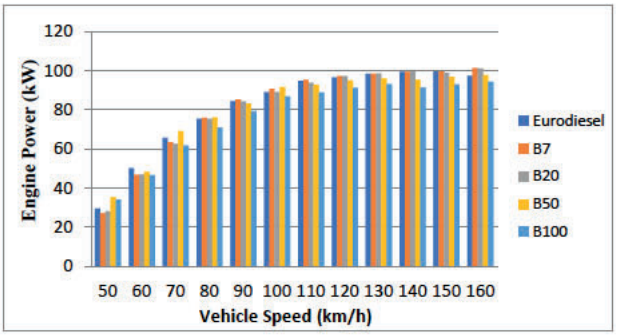


Figure 3. Engine torque change of fuels depending on vehicle speed.

Figure 3 shows the engine power changes depending on the cycle of the test engine in the use of camelina biodiesel fuel mixtures. Maximum engine power values obtained for Euroiesel, B7, B20, B50 and B100 fuels were calculated as 99.6, 101.4, 101.2, 97.7 and 94.3 kW, respectively. Maximum engine power has been measured as 101.4 kW at 160 km / h for B7 fuel. The reason for the decrease in engine power due to the proportion of camelina biodiesel in fuels can be thought to be due to the high density and viscosity of camelina biodiesel.

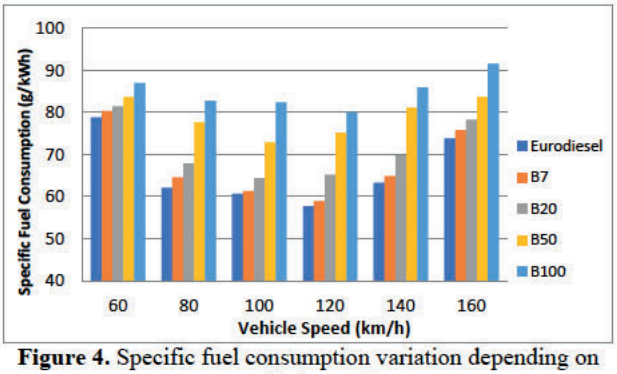


Figure 4. Specific fuel consumption variation depending on vehicle speed

Specific fuel consumption is expressed as the amount of fuel consumed per unit power. Figure 4 shows the effect of biodiesel use on specific fuel consumption. The minimum specific fuel consumption values obtained for Eurodiesel, B7, B20, B50 and B100 fuels were calculated as 57.63, 58.9,64.4, 72.9 and 82.4 g / Kwh, respectively. In the tests conducted, the lowest specific fuel consumption values for all fuels were obtained at a vehicle speed of 120 km / h with eurodiesel, B7 and B100 fuels, while it was obtained at a vehicle speed of 100 km / h with B20 and B50 fuels. In general, specific fuel consumption has increased with the use of biodiesel. Specific fuel consumption increases with the increase in the amount of biodiesel in fuels.

IV. CONCLUSION

In this study, biodiesel was obtained from camelina plant using transesterification method. The effect of camelina biodiesel obtained by using diesel fuel at different rates in an engine with four-cylinder compression ignition common rail fuel system, on engine performance values was investigated. The following results can be obtained by comparing the obtained results with the data obtained with eurodiesel.

Maximum engine power was obtained in B7 fuel with 101.4 kW at 160 km / h. In addition, when compared to 160 km / h, engine power has been obtained in B100, B50, B20 and Eurodiesel fuel mixtures. At the same time, compared to eurodiesel fuel at all engine speeds, the average engine power decreased by 3.28% in the B100 fuel mixture at most. We can say that the reason for the increase of engine power in biodiesel fuel mixes compared to eurodiesel fuel is that the oxygen content in biodiesel contributes to complete combustion.

The maximum engine torque was obtained at 80 km / h with 277 Nm in the B7 fuel mixture. The lowest specific fuel consumption values for all fuels were obtained in Eurodiesel, while the highest specific fuel consumption was obtained in B100. Compared to Eurodiesel fuel, the main reason for the increase of specific fuel consumption in all blends is that a greater amount of fuel is required to achieve the same power.

As a result, the use of biodiesel fuel produced from waste oils and vegetable oils as an alternative fuel without making any changes to vehicles with diesel engines has a positive effect on vehicle performance. For the country's economy, biodiesel fuels need to be diversifying and cultivating products that can produce biodiesel suitable for areas that cannot be cultivated.

ACKNOWLEDGMENT

This research was supported by Selçuk University Scientific Research Projects Coordination Unit with project number 20201022.

REFERENCES

[1] İ. Örs, N. Tarakcıoğlu ve M. Ciniviz, "Yakıt Olarak Benzin–Etanol Karışımlarının Taşıt Performansı ve Egzoz Emisyonlarına Etkisi." Politeknik Dergisi 12.1: 13-19, 2009.

[2] M. Çelik, İ. ÖRS, C. Bayındırlı, M. Demiralp, Experimental investigation of impact of addition of bioethanol in different biodiesels, on performance, combustion and emission characteristics. Journal of Mechanical Science and

Technology, 31(11), 5581-5592., Doi: 10.1007/s12206-017-1052-x, 2017.
[3] M. Çelik, C. Bayındırlı, Enhancement performance and exhaust emissions of rapeseed methylester by using n-hexadecane and n-hexane fuel additives. ENERGY, 202(117643), 1-12., Doi: 10.1016/j.energy.2020.117643, 2020.
[4] J. Zubr, Qualitative variation of Camelina sativa seed from different locations. Industrial Crops and Products, 17(3), 161-169, 2003.
[5] T. Ayaşan, Ketencik Bitkisinin (Camelia sativa) Kanatlı Beslenmesinde Kullanılması. Kahramanmaraş Sütçü İmam Üniversitesi Doğa Bilimleri Dergisi, 17(2), 10-13, 2014.
[6] A.E. Özçelik, H. Aydoğan, M. Acaroğlu, Determining the performance, emission and combustion properties of camelina biodiesel blends, Energy Conversion and Management, Selçuk Üniversitesi, 96.47–57, 2015.
[7] M. Talibi, P. Hellier, M. Watkinson, N. Ladommatos, "Comparative analysis of H_2-diesel co-combustion in a single cylinder engine and a chassis dynamometer vehicle" 44(2): 1239-1252, 2018.
[8] P. Sakthivel, K.A. Subramanian and R. Mathai, "Comparative studies on combustion, performance and emission characteristics of a two-wheeler with gasoline and 30% ethanol-gasoline blend using chassis dynamometer. Applied Thermal Engineering, " 146, 726-737, 2019.
[9] H. Aydoğan, M. Acaroğlu, A.E. Özçelik, Effects Of E2, E20, E30 Bioethanol-Gasoline Blends On Vehicle Emissions and Performance, International Congress on Automotive and Transport Engineering (CONAT 2010), Brasov, Romania, 26-29 October 2010.

Investigation of the Performance of GALILEO-only Precise Point Positioning

O. ATIZ[1], S. ALCAY[1] and S. OGUTCU[1]

[1] Necmettin Erbakan University, Konya/Turkey, oatiz@erbakan.edu.tr
[1] Necmettin Erbakan University, Konya/Turkey, salcay@erbakan.edu.tr
[1] Necmettin Erbakan University, Konya/Turkey, sermetogutcu@erbakan.edu.tr

***Abstract* – In recent years, the number of Global Navigation Satellite Systems (GNSS) reached a total of 4, after launching GALILEO (Europe's Global Navigation Satellite System) and BEIDOU (China's Global Navigation Satellite System). Unlike others, GALILEO is a civilian-based global navigation system and managed by the European Space Agency (ESA). GALILEO will be the fourth satellite-based global positioning service after completing the full operational capability (FOC) mission. First GALILEO satellite launched in 2005 with an experimental test satellite and continued periodically until today. In 2016, GALILEO had 4 in-orbit validation (IOV) satellites and 14 FOC satellites. By the end of 2018, GALILEO is considerably developed with the launch of 8 new FOC satellites and planning to complete the FOC mission in late 2020. Nowadays, 3 IOV satellites and 21 FOC satellites can be used globally. In this study, GALILEO-only precise point positioning (PPP) performance is examined with the latest improvements. For this purpose, the data of 3 IGS (International GNSS Service) stations for a one-week period in 2020 are processed. The obtained coordinates are transformed to the topocentric system (north, east, up), using IGS weekly solutions as the reference. The results are analyzed in terms of accuracy and precision by considering the statistical values (max, min, and RMSE).**

***Keywords* – ESA, GALILEO, GNSS, IGS, PPP**

I. Introduction

Besides GPS and GLONASS, some other constellations at the global scale are under construction. Global Navigation Satellite Systems (GNSS) reached a total number of 4, with adding GALILEO (Europe's GNSS) and BEIDOU (China's GNSS) [1,2]. Among the other GNSS's, GALILEO which is managed by the European Space Agency (ESA), is the first civilian-based service and offering precise positioning and some other civilian services [3]. In the last years, the number of GALILEO satellites drastically increased.

Precise point positioning (PPP) is a cost-effective and powerful tool for obtaining the precise position of the receiver using precise satellite orbit and clock products publishing by analysis centers [4]. The utilization of PPP has increased since it has advantages such as the simplified field experiment, easy of usage, rapid and robust positioning [5]. The main advantage of PPP than differential technique is that it provides a similar accuracy of positioning around the world using a single receiver [6]. Thus, PPP has been widely used in scientific researches for many GNSS applications such as plate tectonic [7], displacement monitoring [8,9] and atmospheric modeling [10]. As of September 2020, GALILEO has enough number of satellites for GALILEO-only PPP. Some studies have been conducted using GALILEO satellites for PPP [11-15]. According to these studies, GALILEO-only PPP has a longer convergence time than GPS or GPS+GLONASS PPP. In terms of accuracy and convergence time, GALILEO contributed to GPS+GLONASS PPP. GALILEO contribution to multi-GNSS is greater at the higher elevation cut-off angles. Also, GALILEO gives more reliable results than GPS in the case of multipath since GALILEO signals have greater SNR (Signal to Noise Ratio) values than GPS. Using precise products produced by different ACs gives similar results except for Wuhan AC. Besides, ambiguity fixing improves accuracy.

Today, 24 of GALILEO satellites can be used, including 2 highly eccentric satellites. The studies above mentioned were mostly investigated the GALILEO contribution of multi-GNSS. Even if some studies are conducted as GALILEO-only, the up-to-date situation can be reviewed. Therefore, the purpose of this study is to investigate the latest performance of GALILEO-only PPP. Thus, the static positioning performance of GALILEO-only PPP examined using GipsyX scientific GNSS data processing software.

This paper is organized as follows. In section 2, the development of GALILEO and the current status is described. In section 3, functional model of PPP is briefly described. Section 3 describes the data and processing strategy. The results are discussed in Section 5, and the conclusions are given in Section 6.

II. Galileo Satellite System

GALILEO is a satellite-based positioning service aiming to serve precise positioning on a global scale and managed by ESA. The GALILEO service consists of three phases, namely, definition, in-orbit-validation, full deployment. After completing the full operational capability (FOC) mission, GALILEO will have some other global scale advanced services [16]. Apart from Open Service (OS) which is the main positioning service, there are some other services such as High accuracy Service (HAS), Public Regulated Service (PRS), and Search and Rescue Service (SAR). The design constellation status of GALILEO consists of 30 Medium Earth Orbit (MEO) satellites. Among these, 24 satellites are operational, and 6

satellites are spares. The satellites are planned to be at 23,200 km above the Earth in three orbital planes with an inclination angle of 55 degree.

The first experimental satellite of GALILEO, the so-called GIOVE-A was launched in 2005. The main purpose of the GIOVE-A satellite was in-orbit testing of vital technologies such as atomic clocks and new navigation signals. GIOVE-A was equipped with a rubidium atomic clock. Also, allocations of frequencies given by the International Telecommunications Union (ITU) were secured. After that GIOVE-B satellite which is equipped with a passive hydrogen maser (PHM) atomic clock -the main atomic clock of GALILEO system- was launched in 2008 [17]. The illustrations of GIOVE-A and GIOVE-B satellites are shown in Figure 1.

Figure 1: GIOVE-A satellite (left) and GIOVE-B satellite (right) [17].

Between 2011 and 2012 years four in-orbit validation (IOV) satellites were placed in orbit. The purpose of IOV satellites was to test and check the system both in space and on Earth [18]. In the last years, GALILEO had significant improvements in the progress of completing the FOC mission. In December 2016, the system achieved to Early Operational Capability (EOC). After that, 14 FOC satellites were placed in orbit until today.

E14 and E18 satellites were placed into highly elliptical orbits than the planned orbit. Even these satellites have highly eccentric orbit, the quality of orbit and its impact on the positioning accuracy is significantly comparable to excluding these satellites [19,20].

As of September 2020, in orbit 24 GALILEO satellite is available [16]. GALILEO is planning to complete its FOC mission in late 2020.

III. Functional Model of PPP

In traditional PPP, due to eliminate the first-order effect of the ionosphere, ionosphere-free combination of code and phase observations are used. Ionosphere-free combination of code and phase equations can be written as below:

$$P^s_{IF,r} = \rho + c * (dt_r - dt^s) + d_{trop} + \epsilon_{P,IF} \quad (1)$$

$$\emptyset^s_{IF,r} = \rho + c * (dt_r - dt^s) + d_{trop} + \lambda_{IF} * N^s_{IF} + \epsilon_{\emptyset,IF} \quad (2)$$

where the superscript s denotes satellite, the superscript r denotes receiver, $P^s_{IF,r}$ and $\emptyset^s_{IF,r}$ are the ionosphere-free combination of code and phase measurement, respectively; ρ is the geometric range between the receiver and satellite, c is the speed of light, dt_r is the receiver clock offset, dt^s is the satellite clock offset, d_{trop} is the tropospheric delay, λ_{IF} is the wavelength of the ionosphere-free combination, N^s_{IF} is the phase integer ambiguity, $\epsilon_{P,IF}$ and $\epsilon_{\emptyset,IF}$ are the measurement noise of ionosphere-free code and phase measurement, respectively. Precise orbit and clock products of one IGS analysis center are used to remove the satellite orbit and clock errors [21]. The hardware delays which are not given in equations above for the sake of simplicity, causes to loss of integer property of ambiguity term [22]. Conventionally, the phase ambiguity term is estimated as float. Other terms which are receiver position, receiver clock, and tropospheric delay are estimated with an adjustment and filtering procedure [23].

IV. Data and Processing Strategy

In this study, the data of 3 IGS (International GNSS Service) stations were used. The geographical distributions of selected stations are given in Figure 1.

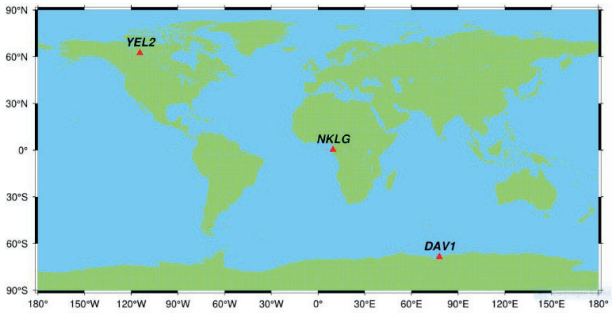


Figure 1: Distribution of selected IGS stations.

Test stations were chosen considering GALILEO satellite visibility and geographical distribution. The receiver and antenna information of the selected stations is given in Table 2. A one-week period (DOY 229 to DOY 235) in 2020 is chosen for static PPP processing. When choosing the stations, due to the reference coordinate necessity, the related IGS weekly solution was checked. Daily data of each station were used as 24-hour sessions with 30 seconds time interval.

Table 2: The receiver and antennas of stations.

Station	Receiver	Antenna and Radome
DAV1	SEPT POLARX5	LEIAR25.R3 + LEIT
NKLG	SEPT POLARX5	TRM59800.00 + SCIS
YEL2	SEPT POLARX5TR	LEIAR25.R4 + NONE

The data processing step was conducted using GipsyX scientific software. GipsyX is a robust GNSS data processing software developed by NASA's Jet Propulsion Laboratory [24]. The detailed processing strategy of GALILEO-only static PPP is given in Table 3.

Table 3: GipsyX data processing strategy.

Parameter	
Adjustment model	Stochastic Kalman filter
Satellite orbit interval	15 min
Satellite clock interval	30 sec
	GALILEO
Frequencies (phase and code)	FREQ1: L1C/L5Q/C1C/C5Q
Epoch interval	30 s
Elevation cut-	10
A priori precisions	Phase: 0.01 m Code: 1 m
A priori troposphere	GPT2 model [25]
Satellite phase center	Corrected by using up-to-date IGS14.atx
Receiver phase center	The PCO and PCV corrections for GPS signals from IGS14.atx are used for GALILEO signals
Ionosphere	First-order effect: Removed by ionosphere-free linear combination. Second-order effect: Removed using JPL's IONEX file.
Zenith wet delay estimation	Random walk 0.5 mm/sqrt(sec)
Horizontal delay gradients estimation	Random walk 0.05 mm/sqrt(sec)
Solid earth tide, ocean tide loading, and polar tides	IERS conventions 2010 [26]
Phase windup	Corrected [27]
Cycle slip	Corrected
Receiver clock jump	Corrected

As a preliminary test of the selected stations, the average, maximum and minimum satellite visibilities for each station and day of year are given in Figure 2-4.

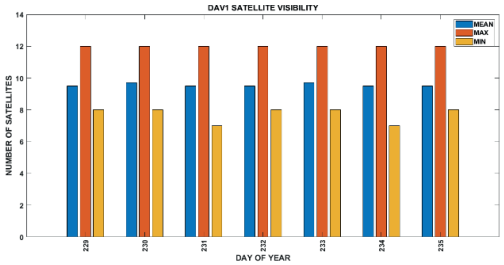


Figure 2: DAV1 average satellite visibility.

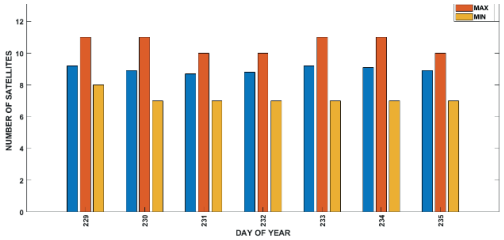


Figure 3: NKLG average satellite visibility.

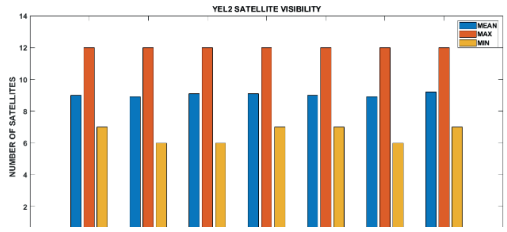


Figure 4: YEL2 average satellite visibility.

According to visibility graphics, there is not a significant difference between the stations. Also, the selected stations can observe an average of 9 GALILEO satellite each day. It can be regarded that the selected stations are suitable for this study.

V. RESULTS

For each day and station cartesian coordinate values were obtained. To analyze the results, cartesian coordinate values were transformed to the topocentric coordinate system which can be expressed as north, east, up components. The transformation was done using IGS weekly solutions as reference. The topocentric coordinate components for each station and day of year are given in Figure 5-7.

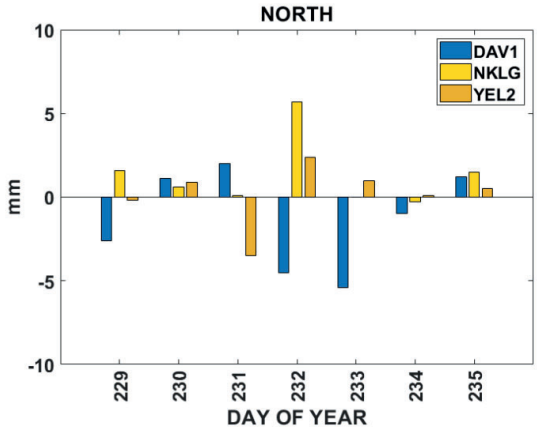


Figure 5: The coordinates for north component.

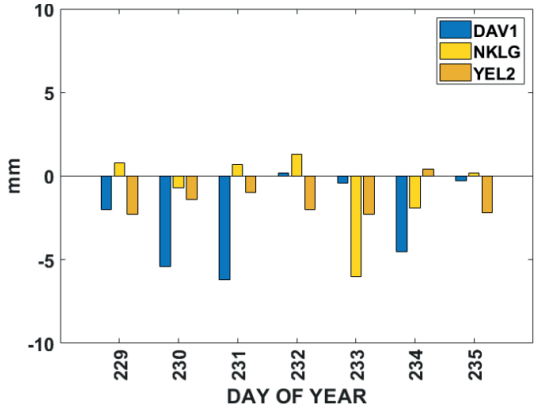


Figure 6: The coordinates for east component.

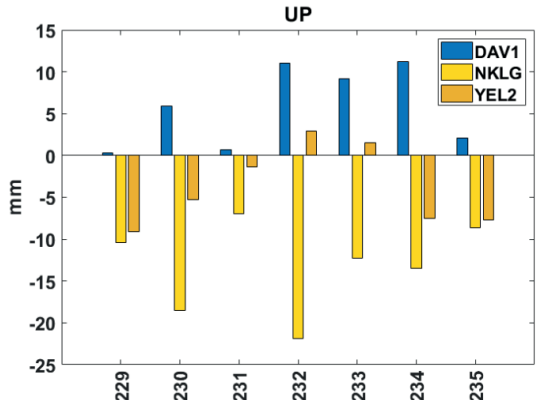


Figure 7: The coordinates for up component.

As seen from the Figures above, GALILEO-only PPP can reach 5 mm, 6 mm, and 20 mm accuracy in the north, east, up components, respectively. The statistical results were calculated for each station from consecutive 7-day results. With these coordinates, root mean square error, std, mean, maximum, and minimum values were examined. The statistical results are given in Table 4.

Table 4: Statistical values for coordinates.

Stations	Parameter	North (mm)	East (mm)	Up (mm)
DAV1	Maximum (abs)	5.4	6.2	11.2
	Minimum (abs)	1.0	0.2	0.3
	Mean	-1.3	-2.7	5.8
	RMSE	3.0	3.6	7.3
	Std	2.7	2.5	4.4
NKLG	Maximum (abs)	5.7	1.9	21.9
	Minimum (abs)	0.0	0.2	7.0
	Mean	1.3	-0.8	-13.2
	RMSE	2.3	2.5	14.1
	Std	1.9	2.3	5.0
YEL2	Maximum (abs)	3.5	2.3	9.1
	Minimum (abs)	0.1	0.4	1.4
	Mean	0.2	-1.5	-3.8
	RMSE	1.7	1.8	5.8
	Std	1.7	0.9	4.4

For the north component, NKLG station has the worst coordinates as it reached a maximum of 5.7 mm. For the east component, DAV1 station has the worst coordinates as it reached a maximum of 6.2 mm. For the up component, the maximum values are 21.9 mm, 11.2 mm and 9.1 mm for NKLG, DAV1 and YEL2 stations, respectively.

The results of both three stations are similar for north and east components. The RMS errors are less than 4 mm. DAV1 station has an absolute maximum value of 6.2 mm on east component. For the up component, as expected, RMS errors are larger than north and east components. RMS errors are 7.3 mm, 14.1 mm, 5.8 mm for DAV1, NKLG and YEL2 stations, respectively.

From the result of the analysis, it is shown that the GALILEO-only PPP method can give mm-cm accuracy in both the vertically and horizontally.

VI. Conclusions

In this study, the latest performance of GALILEO-only PPP is investigated considering satellite visibilities. The results are examined in terms of accuracy and precision. The results showed that GALILEO-only PPP can provide a few millimeters level accuracy for horizontal whereas it can provide a few centimeters level accuracy for vertical. The std values which indicates the precision are 3.0 mm for horizontal component and 5.0 mm for vertical component.

Considering that FOC is not fully completed yet, the results are quite good as other GNSSs. Also, the results are consistent with other studies. For the static positioning with daily GNSS data, GALILEO can successfully be used as of 2020. Furthermore, it is hoped that, after the completing FOC mission, GALILEO will be promising better results with new satellites.

Acknowledgment

The authors would like to thank to NASA Jet Propulsion Laboratory (JPL) for providing the license for GipsyX software. The authors also would like to thank IGS for providing GNSS data and precise products.

References

[1] Guo, F , Li, X , Zhang, X , & Wang, J (2017) Assessment of precise orbit and clock products for Galileo, BeiDou, and QZSS from IGS Multi-GNSS Experiment (MGEX) GPS solutions, 21(1), 279-290 Doi: https://doi org/10 1007/s10291-016-0523-3

[2] Montenbruck, O , Steigenberger, P , Prange, L , Deng, Z , Zhao, Q , Perosanz, F , & Schmid, R (2017) The Multi-GNSS Experiment (MGEX) of the International GNSS Service (IGS)–achievements, prospects, and challenges Advances in space research, 59(7), 1671-1697 Doi: https://doi org/10 1016/j asr 2017 01 011

[3] Li, X , Zhang, X , Ren, X et al Precise positioning with current multi-constellation Global Navigation Satellite Systems: GPS, GLONASS, Galileo and BeiDou Sci Rep 5, 8328 (2015) Doi: https://doi org/10 1038/srep08328

[4] Zumberge, J F , Heflin, M B , Jefferson, D C , Watkins, M M , and Webb, F H , 1997 Precise point positioning for the efficient and robust analysis of GPS data from large networks Journal of geophysical research: solid earth, 102(B3), 5005-5017 Doi: https://doi org/10 1029/96JB03860

[5] Alkan, R M , İlçi, V , Ozulu, I M , & Saka, M H (2015) A comparative study for accuracy assessment of PPP technique using GPS and GLONASS in urban areas Measurement, 69, 1-8 Doi: https://doi org/10 1016/j measurement 2015 03 012

[6] Ogutcu, S , 2020a Performance analysis of ambiguity resolution on PPP and relative positioning techniques: consideration of satellite geometry International Journal of Engineering and Geosciences, 5(2), 73-93 Doi: https://doi org/10 26833/ijeg 580027

[7] Shi, C , Lou, Y D , Zhang, H P , Zhao, Q , Geng, J , Wang, R , & Liu, J (2010) Seismic deformation of the Mw 8 0 Wenchuan earthquake from high-rate GPS observations Advances in Space Research, 46(2), 228-235 Doi: https://doi org/10 1016/j asr 2010 03 006

[8] Alcay, S , Ogutcu, S , Kalayci, I , & Yigit, C O (2019) Displacement monitoring performance of relative positioning and Precise Point Positioning (PPP) methods using simulation apparatus Advances in Space Research, 63(5), 1697-1707 Doi: https://doi org/10 1016/j asr 2018 11 003

[9] Martín, A , Anquela, A B , Dimas-Pagés, A , & Cos-Gayón, F (2015) Validation of performance of real-time kinematic PPP A possible tool for

deformation monitoring Measurement, 69, 95-108 Doi: https://doi org/10 1016/j measurement 2015 03 026
[10] Pikridas, C , Katsougiannopoulos, S , & Zinas, N (2014) A comparative study of zenith tropospheric delay and precipitable water vapor estimates using scientific GPS processing software and web based automated PPP service Acta Geodaetica et Geophysica, 49(2), 177-188 Doi: https://doi org/10 1007/s40328-014-0047-7
[11] Xia, F , Ye, S , Xia, P , Zhao, L , Jiang, N , Chen, D , & Hu, G (2019) Assessing the latest performance of Galileo-only PPP and the contribution of Galileo to Multi-GNSS PPP Advances in space research, 63(9), 2784-2795 Doi: https://doi org/10 1016/j asr 2018 06 008
[12] Hadas, T , Kazmierski, K , & Sośnica, K (2019) Performance of Galileo-only dual-frequency absolute positioning using the fully serviceable Galileo constellation GPS Solutions, 23(4), 108 Doi: https://doi org/10 1007/s10291-019-0900-9
[13] Ogutcu, S (2020) Assessing the contribution of Galileo to GPS+ GLONASS PPP: Towards full operational capability Measurement, 151, 107143 Doi: https://doi org/10 1016/j measurement 2019 107143
[14] Afifi, A , & El-Rabbany, A (2014) Enhanced model for precise point positioning with single and dual frequency GPS/GALILEO observables International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences Doi: https://doi org/10 5194/isprsarchives-XL-2-173-2014
[15] Liu, G , Zhang, X , & Li, P (2019) Improving the Performance of Galileo Uncombined Precise Point Positioning Ambiguity Resolution Using Triple-Frequency Observations Remote Sensing, 11(3), 341 Doi: https://doi org/10 3390/rs11030341
[16] The current constellation of the Galileo system https://www gsc-europa eu/system-service-status/constellation-information Accessed: 09 10 2020
[17] ESA Applications, Navigation, http://www esa int/Applications/Navigation Accessed: 09 10 2020
[18] Kiliszek, D , & Kroszczyński, K (2020) Performance of the precise point positioning method along with the development of GPS, GLONASS and Galileo systems Measurement, 108009 Doi: https://doi org/10 1016/j measurement 2020 108009
[19] Sośnica, K , Prange, L , Kaźmierski, K , Bury, G , Drożdżewski, M , Zajdel, R , & Hadas, T (2018) Validation of Galileo orbits using SLR with a focus on satellites launched into incorrect orbital planes Journal of geodesy, 92(2), 131-148 Doi: https://doi org/10 1007/s00190-017-1050-x
[20] Bahadur, B , & Nohutcu, M (2020) Galileo-based precise point positioning with different MGEX products Measurement Science and Technology Doi: https://doi org/10 1088/1361-6501/ab8b85
[21] Kouba, J , and Héroux, P , 2001 Precise point positioning using IGS orbit and clock products GPS solutions, 5(2), 12-28 Doi: https://doi org/10 1007/PL00012883
[22] Xiao, G , Li, P , Sui, L , Heck, B , & Schuh, H (2019) Estimating and assessing Galileo satellite fractional cycle bias for PPP ambiguity resolution GPS Solutions, 23(1), 3 Doi: https://doi org/10 1007/s10291-018-0793-z
[23] ESA Navipedia, Precise Point Positioning https://gssc esa int/navipedia/index php/Precise_Point_Positioning Accessed: 29 09 2020
[24] Bertiger, W , Bar-Sever, Y , Dorsey, A , Haines, B , Harvey, N , Hemberger, D , & Murphy, D (2020) GipsyX/RTGx, A New Tool Set for Space Geodetic Operations and Research Advances in Space Research, 66(3), 469-489 Doi: https://doi org/10 1016/j asr 2020 04 015
[25] Lagler, K , Schindelegger, M , Böhm, J , Krásná, H , & Nilsson, T (2013) GPT2: Empirical slant delay model for radio space geodetic techniques Geophysical research letters, 40(6), 1069-1073 Doi: https://doi org/10 1002/grl 50288
[26] Petit, G , Luzum, B , "IERS Conventions", Bureau International des Poids et Mesures Sevres (France), 2010 Available online: https://www iers org/IERS/EN/Publications/TechnicalNotes/tn36 html Accessed: 11 10 2020
[27] Wu, J T , Wu, S C , Hajj, G A , Bertiger, W I , & Lichten, S M (1992) Effects of antenna orientation on GPS carrier phase asdy, 1647-1660

The Kinematic Results of the First Open-Source Precise Point Positioning-Ambiguity Resolution Software

O. ATIZ[1], S. OGUTCU[1], S. ALCAY[1] and I. KALAYCI[1]

[1] Necmettin Erbakan University, Konya/Turkey, oatiz@erbakan.edu.tr
[1] Necmettin Erbakan University, Konya/Turkey, sermetogutcu@erbakan.edu.tr
[1] Necmettin Erbakan University, Konya/Turkey, salcay@erbakan.edu.tr
[1] Necmettin Erbakan University, Konya/Turkey, ikalayci@erbakan.edu.tr

***Abstract* – Today, traditional precise point positioning (PPP) is an essential technique for most scientific studies such as geodesy and plate tectonics. In recent years, the scientific community is mostly concerned about PPP with the ambiguity resolution (PPP-AR) approach. Hereby, Analysis Centers (ACs) started to produce phase bias products which are necessary for ambiguity resolved PPP. Furthermore, different software packages have been developed to apply AR (ambiguity resolution). PRIDE-PPPAR software is the first open-source and free PPP-AR software, developed by Wuhan University GNSS Research Center. The software works with CODE's (Center of Orbit Determination) precise ephemeris and clock corrections, suitable with WHU's (Wuhan Analysis Center) phase bias products. In this study, the kinematic PPP performance of PRIDE-PPPAR software is examined. Daily data of 5 IGS (International GNSS Service) stations for a one-week period in 2020 processed under different (7° and 30°) cut-off angles. Coordinate values are estimated using two PPP modes (ambiguity-float and ambiguity-fixed) to find out the contribution of AR. The results are examined regarding to the accuracy and precision. Also, wrong fixing percentages are calculated to investigate the success of AR.**

***Keywords* – AR, GNSS, IGS, PPP, PRIDE-PPPAR.**

I. Introduction

PRECISE point positioning (PPP) is an absolute positioning method that can give centimeter-level positioning accuracy after a convergence time using precise satellite orbit and clock products [1,2]. Traditional PPP has become a practical tool since the reference station requirement is eliminated. There are many studies conducted using PPP such as structural health monitoring and observing seismic waves. [3-5]. The main drawback of PPP technique is that it needs a long convergence time to achieve the sub-decimeter level positioning accuracy. With the increasing number of GNSS constellations, the multi-GNSS concept is come up [6]. The multi-GNSS PPP is emerged by supplementing GLONASS, Galileo, and Beidou constellations to PPP model. With this development, convergence time has shortened, and accuracies are improved [7,8].

The high interest in PPP has accelerated the improvements of the technique. Conventionally, PPP method offers an ambiguity-float solution due to the difficulties of ambiguity resolution (AR). In the last years, new approaches have been developed for ambiguity-fixed PPP or called PPP-AR (PPP with ambiguity resolution) [9-11]. Mainly, the uncalibrated hardware delays (UHD) arising in both receivers and satellites obstructs the AR [15]. To overcome this deficiency additional phase-clock bias products are needed, which are provided by IGS analysis centers (ACs), hereafter called AR products. [16]. However, the AR products can only be used with the consistent precise satellite orbit and clock products.

Some studies have been conducted using PPP-AR approach. Li, et al. in 2018 [12] investigated the impact of AR on accuracy. The results show that AR was improved the positioning accuracy. Goudarzi and Banville in 2018 [13], compared the PPP-AR and relative method for detecting surface deformations. According to the results, PPP-AR can be used successfully to detect movements in the long observation period. Hu et al. in 2014 [14], used PPP-AR for monitoring land deformations in a mining area. They indicated that AR was shortened the convergence time.

In addition, different software packages have been developed to apply AR. Among them, PRIDE-PPPAR is a new open-source software developed by PRIDE Lab at Wuhan University [17]. PRIDE-PPPAR works with CODE (Center of Orbit Determination) AC's final orbit and clock corrections, and AR products of WHU (Wuhan University) AC as default. Unlike the others, PRIDE-PPPAR is designed especially for PPP-AR with an enhanced AR algorithm. Also, PRIDE-PPPAR is the first open-source PPP-AR software. Thus, it is worth to investigate the positioning performance of this software.

In this study, the kinematic positioning performance of first open-source PPP-AR is investigated considering different (7° and 30°) cut-off angles. Due to the inability of processing other constellations in PRIDE-PPPAR software, only GPS satellites are considered. The data of 5 IGS (International GNSS Service) stations for a one-week period in 2020 processed using two PPP modes (ambiguity-fixed and ambiguity-float). The contribution of AR to PPP is examined considering accuracy and precision.

This paper is organized as follows. In section 2, the functional model of PPP is shortly introduced. Section 3 describes the data and processing strategy. The results are

discussed in Section 4, and the conclusions are given in Section 5.

II. Ionosphere-Free PPP Functional Model

In general, ionosphere-free (IF) combinations of code and phase observations are used for PPP, to eliminate the first-order ionosphere effect. The observation equations can be written as:

$$P^s_{IF,r} = \rho + c * (dt_r - dt^s) + d_{trop} + HD_{P,r,IF} - HD^{P,s,IF} + \epsilon_{P,IF} \quad (1)$$

$$\emptyset^s_{IF,r} = \rho + c * (dt_r - dt^s) + d_{trop} + \lambda_{IF} * (N^s_{IF} + HD_{\emptyset,IF,r} - HD^{\emptyset,IF,s}) + \epsilon_{\emptyset,IF} \quad (2)$$

where the superscript r and superscript s denotes receiver and satellite, respectively; $P^s_{IF,r}$ and $\emptyset^s_{IF,r}$ are the IF combination of code and phase measurement, respectively; ρ is the geometric range, c is the speed of light, dt_r is the receiver clock offset, dt^s is the satellite clock offset, d_{trop} is the tropospheric delay, λ_{IF} is the wavelength of the IF combination, N^s_{IF} is the phase integer ambiguity, $HD_{P,r,IF}$ and $HD_{\emptyset,IF,r}$ indicates the receiver hardware code and phase delays, respectively; $HD^{P,s,IF}$ and $HD^{\emptyset,IF,s}$ indicates the satellite hardware code and phase delays, respectively; $\epsilon_{P,IF}$ and $\epsilon_{\emptyset,IF}$ are the measurement noise of IF code and phase measurement, respectively.

In general, IGS precise orbit and clock products are used to eliminate the clock offsets. The satellite hardware code bias is lumped into satellite clock offset term since IGS clock products are computed using IF code measurements. Similarly, receiver hardware code bias can be absorbed by the receiver clock offset term. The receiver and satellite hardware phase biases are lumped into the ambiguity parameter. Therefore, the integer nature of the ambiguity term is destroyed [18].

The IF code and phase observables can be written as:

$$P^s_{IF,r} = (f1^2 * P1^s - f2^2 * P2^s)/(f1^2 - f2^2) \quad (3)$$

$$\emptyset^s_{IF,r} = (f1^2 * \emptyset1^s - f2^2 * \emptyset2^s)/(f1^2 - f2^2) \quad (4)$$

where $f1$ and $f2$ two frequencies in Hertz; $P1$, $P2$, $\emptyset1$, and $\emptyset2$ are the code-pseudorange and carrier-phase observations on two frequencies. IF wavelength and ambiguity parameters can be written as:

$$\lambda_{IF} = \frac{f1^2}{f1^2 - f2^2} * \lambda_1 - \frac{f2^2}{f1^2 - f2^2} * \lambda_2 \quad (5)$$

$$\widetilde{N_{IF}} = \frac{f1^2 * N_1}{f1^2 - f2^2} - \frac{f2^2 * N_2}{f1^2 - f2^2} \quad (6)$$

As can be seen from Equation 2 to estimate the integer value of IF ambiguity, satellite and receiver hardware phase bias information is needed. As mentioned before ACs are computing satellite hardware biases from the network and providing to users, so-called AR products. Herein, receiver dependent hardware biases can be safely ignored for PPP [19].

In general, estimating the ambiguity-fixed solution is done in three sequential steps. Firstly, the ambiguity-float solution is estimated without considering the integer ambiguity term. Secondly, the float solution is used to solve the integer value of the ambiguity term. Then, the unknown parameters are calculated using the integer value of ambiguity [20].

IF ambiguity can be estimated by wide-lane (WL) and narrow-lane (NL) combinations [21].

$$\emptyset_{WL} = (f1 * \emptyset1 - f2 * \emptyset2)/(f1 - f2) \quad (7)$$

$$\emptyset_{NL} = (f1 * \emptyset1 + f2 * \emptyset2)/(f1 + f2) \quad (8)$$

$$P_{WL} = (f1 * P1 - f2 * P2)/(f1 - f2) \quad (9)$$

$$P_{NL} = (f1 * P1 + f2 * P2)/(f1 + f2) \quad (10)$$

where $\emptyset_{WL}$ and P_{WL} are WL combinations of phase and code, respectively; $\emptyset_{NL}$ and P_{NL} are NL combinations of phase and code, respectively. The WL and NL ambiguities cannot be calculated and fixed simultaneously, since NL ambiguity contains WL ambiguity. Therefore, first WL ambiguity is resolved then NL ambiguity is resolved. IF ambiguity can be fixed only both WL and NL ambiguities are fixed [22]. When NL ambiguities cannot be fixed to an integer, AR can be conducted as partial fixing [23].

III. Data and Processing Strategy

Five globally distributed IGS stations were chosen for data processing. In the processing, daily data of a one-week period (DOY 12-18) in 2020 is used. Figure 1 shows the geographical distribution of selected IGS stations. When choosing stations, it was checked that they were in related IGS weekly solutions, since IGS solutions are used as reference coordinates. The data availability for RINEX files was checked and the files over %98 availability were used for processes.

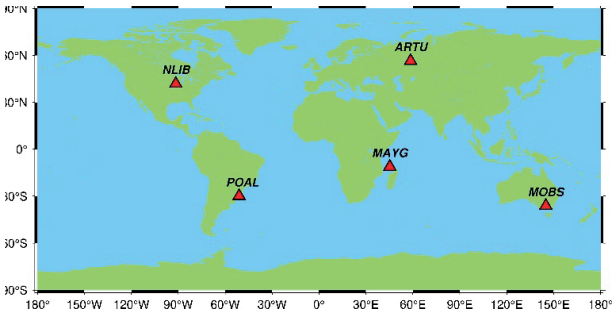


Figure 1: Distribution of selected IGS stations.

Only GPS data was used since in the latest version of PRIDE-PPPAR other GNSSs are not supported. The positioning mode was set to kinematic. Satellite elevation cut-off angles were set to 7° and 30°, to represent both open sky and urban canyon conditions. Processes were conducted in two different PPP modes, namely, ambiguity-float and ambiguity fixed. In ambiguity resolution, a bootstrapping method was used for both WL and NL ambiguities. The detailed PPP processing parameters are given in Table 1.

Table 1: The detailed processing parameters.

	Kinematic
Satellite ephemeris and clock	CODE final orbit/ Estimated
Observation epoch interval	30 s
Filter	Forward and backward
Cut-	7 /30
	Applied
Satellite phase center	Corrected with up-to-date IGS14.atx
Receiver phase center	Corrected with up-to-date IGS14.atx
Ionospheric effect	Eliminated by IF linear combination
Phase ambiguities	FLOAT/FIXED
Ambiguity resolution	Bootstrapping method is used for both WL and NL
Intra-Frequency Bias	GPS C1-P1 code bias was corrected with up-do-date DCB file
Coordinate estimation	White-noise-like parameters
Troposphere	GMF model with piece-wise constant with Saastamoinen model
Zenith wet delay estimation	Piece-wise constant
Horizontal delay gradients estimation	Piece-wise constant
Phase wind-up	Corrected
Tidal effects	Solid tide, ocean loading and polar tides (IERS 2003)

IV. Results and Analysis

The IGS weekly combined solutions were assumed as true coordinates. With these coordinates, the obtained solutions were converted to the topocentric coordinate system (north, east, up). For proper analysis, an outlier test was run with a 10 cm threshold. The coordinates above this were discarded from the root mean square error (RMSE) calculations. For simplicity of visualization, RMS errors of ambiguity-float PPP results for each station and cut-off angle are provided in Table 2.

Table 2: RMSEs calculated from ambiguity-float PPP.

Station	Cut-off: 7		
	North (cm)	East (cm)	Up (cm)
ARTU	0.80	0.87	1.61
MAYG	1.02	1.36	5.08
MOBS	0.87	1.20	2.81
NLIB	1.25	1.35	3.64
POAL	1.02	0.97	2.75
	Cut-off: 30		
	North (cm)	East (cm)	Up (cm)
ARTU	1.48	2.31	4.42
MAYG	1.80	2.17	5.06
MOBS	2.40	2.63	4.89
NLIB	1.58	2.12	5.47
POAL	1.74	2.30	4.68

To compare the results belonging to the same station, after the removal of outliers, an alignment was performed between two PPP modes. For better comprehension, the outlier percentages for two different cut-off angles are given in Figure 2 and Figure 3.

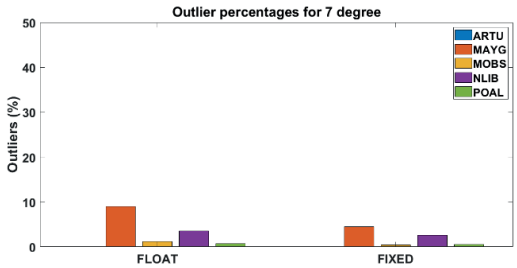


Figure 2: Outlier percentages for 7 cut-off angle.

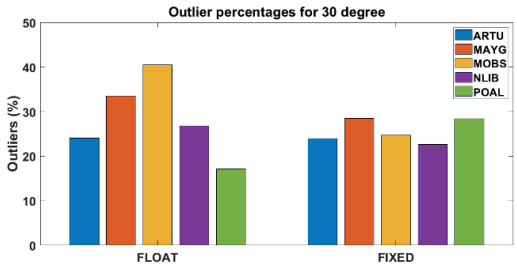


Figure 3: Outlier percentages for 30 cut-off angle.

According to the Figures above, for both two different cut-off angles, the ambiguity-float solutions have more outliers than the ambiguity-fixed solutions. As can be seen from the Figures, the outlier percentages for the 30° cut-off angle are much bigger than the 7° cut-off angle due to the limited skyview of GPS satellites.

In order to investigate the performance of AR, accuracy improvements in north, east and up components from the ambiguity-fixed solutions with respect to the ambiguity-float solutions for each station are given in Figure 4-6.

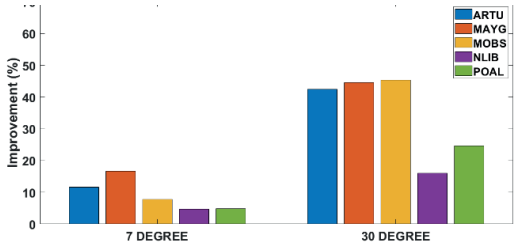


Figure 4: Accuracy improvements of AR in north component.

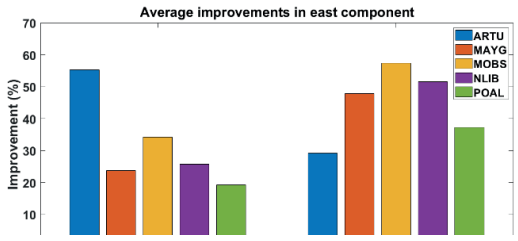


Figure 5: Accuracy improvements of AR in east component.

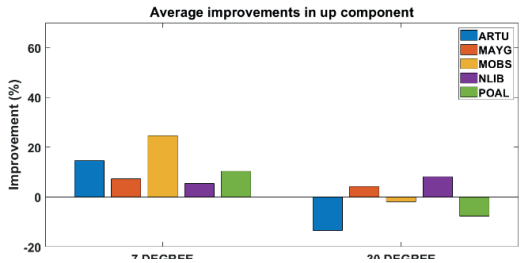


Figure 6: Accuracy improvements of AR in up component.

As can be seen from improvement graphics, ambiguity fixing provides more accurate coordinates. As might be expected, the improvements in the east component are higher than the north and up components for both two cut-off angles. The NLIB station has the lowest improvement in north components. The lowest improvement for the east component is at POAL station. For the up component, the improvements are quite lower at the high cut-off angle. Even ARTU, MOBS and POAL stations are negatively affected by AR. Also, it can be concluded that the improvements for higher cut-off angles are much bigger than lower cut-off angles. The degradation in up component for 30° elevation cut-off angle, can be due to the limited number of satellites.

In addition, the wrong fixing of integer ambiguities was calculated to investigate the performance of AR. If any ambiguity-fixed solution is in the outlier range and its ambiguity-float counterpart is not in the outlier range it so-called wrong fixing. Figure 7 shows the wrong fixing percentages for each station and cut-off angle.

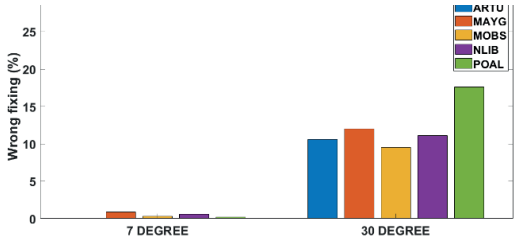


Figure 7: Average wrong fixing percentages.

As can be seen from Figure 7, the wrong fixing of integer ambiguity is not significant for lower elevation cut-off angles. Also, it can be concluded that as the satellite visibility decreases, the success of AR is decreased. The short arc length of satellites under limited skyview leads to unstable AR.

The PRIDE-PPPAR software uses the forward and backward filtering as default. And this cannot be changed to forward filtering. Due to this, the convergence time analysis is not included in this work.

V. Conclusion

In this paper, the kinematic positioning performance of the first open-source PPPAR software (PRIDE-PPPAR) is examined for different elevation cut-off angles (7° and 30°) which are representing the open sky and urban canyon situations. According to the results, ambiguity-fixed solutions are quite better than ambiguity-float solutions. Even at the high elevation cut-off angles, accuracy improvements were found comparable. Also, wrong fixing percentages of the integer ambiguities were investigated. It is found that as satellite visibility decreases, the success of AR is decreased.

In general, PRIDE-PPPAR software has successfully applied the AR. As expected, accuracies for the north, east and up components are improved with the AR. Due to the restricted filtering option of the software, the convergence time analysis could not perform.

Based on this study, PRIDE-PPPAR software can be used as alternative software for PPP-AR.

Acknowledgment

The authors would like to thank to Wuhan University GNSS Center for providing PRIDE-PPPAR software and phase clock/bias products.

References

[1] Zumberge, J F , Heflin, M B , Jefferson, D C , Watkins, M M , & Webb, F H (1997) Precise point positioning for the efficient and robust analysis of GPS data from large networks Journal of geophysical research: solid earth, 102(B3), 5005-5017 Doi: https://doi org/10 1029/96JB03860

[2] Kouba, J , & Héroux, P (2001) Precise point positioning using IGS orbit and clock products GPS solutions, 5(2), 12-28 Doi: https://doi org/10 1007/PL00012883

[3] Leandro, R F , Santos, M C , & Langley, R B (2011) Analyzing GNSS data in precise point positioning software GPS solutions, 15(1), 1-13 Doi: https://doi org/10 1007/s10291-010-0173-9

[4] Kouba, J , 2003 Measuring seismic waves induced by large earthquakes with GPS, Studia Geophysica et Geodaetica, 47 (4), 741–755 Doi: https://doi org/10 1023/A:1026390618355

[5] Yigit, C O , 2016 Experimental assessment of post-processed kinematic Precise Point Positioning method for structural health monitoring Geomatics, Natural Hazards and Risk, 7(1), 360-383 Doi: https://doi org/10 1080/19475705 2014 917724

[6] An, X , Meng, X , & Jiang, W (2020) Multi-constellation GNSS precise point positioning with multi-frequency raw observations and dual-frequency observations of ionospheric-free linear combination Satellite Navigation, 1(1), 1-13 Doi: https://doi org/10 1186/s43020-020-0009-x

[7] Bahadur, B , & Nohutcu, M (2018) PPPH: a MATLAB-based software for multi-GNSS precise point positioning analysis GPS solutions, 22(4), 113 Doi: https://doi org/10 1007/s10291-018-0777-z

[8] Cai, C , Gao, Y , Pan, L , & Zhu, J (2015) Precise point positioning with quad-constellations: GPS, BeiDou, GLONASS and Galileo Advances in space research, 56(1), 133-143 Doi: https://doi org/10 1016/j asr 2015 04 001

[9] Laurichesse, D , Mercier, F , Berthias, J P , Broca, P , & Cerri, L (2009) Integer ambiguity resolution on undifferenced GPS phase measurements and its application to PPP and satellite precise orbit determination Navigation, 56(2), 135-149 Doi: https://doi org/10 1002/j 2161-4296 2009 tb01750 x
[10] Li, X , & Zhang, X (2012) Improving the estimation of uncalibrated fractional phase offsets for PPP ambiguity resolution The Journal of Navigation, 65(3), 513-529 Doi: https://doi org/10 1017/S0373463312000112
[11] Collins, P , Lahaye, F , Heroux, P , & Bisnath, S (2008) Precise point positioning with ambiguity resolution using the decoupled clock model In Proceedings of the 21st international technical meeting of the satellite division of the Institute of Navigation (ION GNSS 2008) (pp 1315-1322)
[12] Geng, J , Teferle, F N , Shi, C , Meng, X , Dodson, A H and Liu,J (2009) Ambiguity resolution in precise point positioning with hourly data GPS solutions, 13(4), 263-270 Doi: https://doi org/10 1007/s10291-009-0119-2
[13] Banville, S , Geng, J , Loyer, S , Schaer, S , Springer, T , & Strasser, S (2020) On the interoperability of IGS products for precise point positioning with ambiguity resolution Journal of geodesy, 94(1), 10 Doi: https://doi org/10 1007/s00190-019-01335-w
[14] Li, X , Li, X , Yuan, Y , Zhang, K , Zhang, X , & Wickert, J (2018) Multi-GNSS phase delay estimation and PPP ambiguity resolution: GPS, BDS, GLONASS, Galileo Journal of geodesy, 92(6), 579-608 Doi: https://doi org/10 1007/s00190-017-1081-3
[15] Goudarzi, M A , & Banville, S (2018) Application of PPP with ambiguity resolution in earth surface deformation studies: a case study in eastern Canada Survey Review, 50(363), 531-544 Doi: https://doi org/10 1080/00396265 2017 1337951
[16] Hu, H , Gao, J , & Yao, Y (2014) Land deformation monitoring in mining area with PPP-AR International journal of mining science and technology, 24(2), 207-212 Doi: https://doi org/10 1016/j ijmst 2014 01 011
[17] Geng, J , Chen, X , Pan, Y , Mao, S , Li, C , Zhou, J , & Zhang, K (2019) PRIDE PPP-AR: an open-source software for GPS PPP ambiguity resolution GPS Solutions, 23(4), 91 Doi: https://doi org/10 1007/s10291-019-0888-1
[18] Xiao, G , Li, P , Gao, Y , & Heck, B (2019) A unified model for multi-frequency PPP ambiguity resolution and test results with Galileo and BeiDou triple-frequency observations Remote sensing, 11(2), 116 Doi: https://doi org/10 3390/rs11020116
[19] Li, P , Zhang, X , Ren, X , Zuo, X , & Pan, Y (2016) Generating GPS satellite fractional cycle bias for ambiguity-fixed precise point positioning GPS solutions, 20(4), 771-782 Doi: https://doi org/10 1007/s10291-015-0483-z
[20] Joosten, P (2000) Fixing the ambiguities-are you sure they're right? GPS world, 11(5), 46-51
[21] Teunissen, P , & Montenbruck, O (Eds) (2017) Springer handbook of global navigation satellite systems Springer
[22] Ge, M , Gendt, G , Rothacher, M A , Shi, C , & Liu, J (2008) Resolution of GPS carrier-phase ambiguities in precise point positioning (PPP) with daily observations Journal of geodesy, 82(7), 389-399 Doi: https://doi org/10 1007/s00190-007-0187-4
[23] Li, P , & Zhang, X (2015) Precise point positioning with partial ambiguity fixing Sensors, 15(6), 13627-13643 Doi: https://doi org/10 3390/s150613627

The Importance of Landscape in the Coronavirus (Covid 19) Period

G.SANDAL ERZURUMLU

Niğde Ömer Halisdemir University, Niğde/Turkey, gpeyzaj@gmail.com

Abstract - **Along with the appearance of Covid-19 outbreak in Turkey since March, people avoided the crowds and arrival of various travel bans led people to houses with gardens. Especially families with children have come to prefer calm areas with beautiful landscapes to spend free time outdoors. In this period, people started to use ornamental plants in hobby gardens. In addition to these, the demands for rural areas with beautiful landscapes have intensified day by day. The aim of this study is to determine the importance given to green areas and landscape in any epidemic period. It is aimed to determine the place of the green areas in human life.**

Keywords: Covid 19, Coronavirus, Detached House, Niğde, Landscape

I INTRODUCTION

Historical information shows that people chose forests and wooded areas in order to live collectively even in the early periods. This is mainly because people can only meet their food, clothing and shelter needs in forest areas. In agriculture (plant and animal husbandry), forests have been used to a great extent in order to reach the current situation. Most of the building materials required for the establishment of villages, towns and large cities were obtained from forests [1]. With the increase in population, residential areas have turned into cities.

Cities are a collection of rapidly developing and changing systems, consisting of many spaces and the relationship between these spaces. Cities are living beings that bring people together in large communities and provide services in many areas such as shelter, protection, education, work, trade, social and cultural functions, living with their residents, which are constantly evolving, changing and growing. Cities are built by people, but they develop and grow on their own [2].

As the population increases in cities, due to the need for residential areas, destruction of nature and zoning areas increase regardless of the fertility level of the soil. As in the current situation, people's desire for nature and green begins to increase. People who want a green area or a small piece of nature (wooded areas) live in quiet and green places.

It is possible to say that human contact with nature can occur in three ways. The first is watching nature from a window, looking at the picture or photo from the book or video [3, 4, 5, 6, 7, 8]. Secondly, doing activities such as sitting on a bench outdoors, reading a book, communicating with people, riding a bike [9, 10]. The third is active engagement level interactions such as gardening, agricultural work, mountain hiking and camping, horseback riding or other physical exercise that cooperate with nature [11, 12].

Human physiology of human interaction with nature [4, 13] and its psychology, (Kaplan & Kaplan 1989, Leather et al. 1998, Kuo 2001, Kuo & Sullivan 2001, Elings 2006) there are many studies that draw attention to the positive effects of these subjects. These studies point out that the use of nature-view video, photo presentations and simulations and nature sound sources with different scales in the interior has positive effects on human psychology in the context of auditory and visual perception levels [12].

Unhealthy conditions in cities have become increasingly threatening to urban life. Unhealthy conditions have caused many diseases to become epidemics in industrial cities, causing frequent mass deaths in cities. However, epidemic diseases in industrial cities pose a threat not only to the working people, but also to the bourgeoisie due to the nature of environmental problems. As a matter of fact, these epidemics have now become a problem in cities [14].

In big cities and industrial settlements, green areas and masses undertake the biggest contribution to the health of people or to regain their health. This contribution cannot be denied, especially in arid climatic regions, in regions where natural vegetation has been largely or completely destroyed. In a forested area, picnic, camping, even a few hours of strolling and resting in various formations of natural vegetation affect human health in a large and positive way. It is also possible to protect wildlife from environmental harmful factors most effectively with green areas.

II RESEARCH FINDINGS

A Respiratory Diseases From Past To Present

There have been many epidemics in the world from past to present. When the diseases related to respiration that cause distressing destructions were examined, the following were determined.

- While small and large-scale epidemics have claimed more than 110 million lives in various parts of the world since the early 1900s, the biggest epidemics in history are Black Death and Spanish flu. Considered one of the biggest epidemics in the history of the world, the Spanish flu had its effect for 18 months between 1918 and 1920, causing the death of nearly 50 million people worldwide.
- It is estimated that 200 thousand people died in the world when the disease, which is described by scientists as a mixture of influenza in humans and influenza in pigs and birds, caused an epidemic in 2009.
- The world faced the global epidemic once again with a new disease that broke out in Wuhan, China in the last days of 2019. Contagious Covid-19, spread rapidly around the world and has caused the death of more than 1 million people to date.

As can be seen, respiratory diseases have occurred in history and today have caused the death of many people. Especially in these periods, people's demand for nature for rest and psychological relaxation has increased. In the period of epidemics, people need more landscape, nature and forest

areas. Ecology degradation affects the survival of many species. The continuity of ecological balance is required for the continuity of life. There are important concepts to landscape ecology. Biodiversity is the relationship between the ecosystem and the function between the two, the continuity between spatially heterogeneous connections (animate-non-living relationship), disturbance, resistance and recovery. Ecosystem, ecology and landscape begin with nature conservation. To achieve this, protecting the ecology and ensuring the sustainability of the landscape should be considered first.

B Ecology, Urban Ecosystem and the Impact of Sustainable Landscape on People

Ecology is the science that studies the relationships of living beings with each other and their environment. Ecology, integrated with technology, is reflected in the space and creates an alternative city vision in harmony with nature, economically efficient and socially compatible. Ecological life, perceived as an alternative way of life, should become a natural part of everyone's lifestyle over time.

Urban ecosystems are places where human influence is seen due to residential, industrial and commercial use. They also have distinctive ecological features and accommodate many different habitats. Urban ecosystems are under threat, especially with industrialization and urbanization [15]. The benefits of the urban ecosystem are to conserve and enrich biological diversity, create habitats for living things, soil formation, seed transport, pollination and contribute to the life cycle.

Sustainable landscape is defined as the process of improving, preserving and developing all environmental values that constitute the environment of both today and future generations without endangering the existence and quality of resources that future generations will need. University campuses, which have an important place in the urban fabric, are one of the environments where environmental problems occur on a micro scale [16]. These definitions that make up the landscape will contribute to the protection and continuity of nature. Landscape areas will provide opportunities for recreation, increase of environmental quality, better physical and mental health of urban people and nature education and related activities. These criteria, which are important for human health, will be needed more especially during pandemic periods. In the period of Covid 19, the fact that people had to stay away from the crowd, the closing of entertainment places increased the demands for green spaces and houses with gardens. While spaces for hobby gardens were reserved in houses with gardens, interest in ornamental plants increased. While Human-nature connection is considered as a potential remedy, nature is harmed by humans. Landscape emerges as a product of nature and human and its culture in varying proportions. In different parts of the world where climate and soil conditions are almost the same, the cultural levels, traditions, social and economic conditions of people who use it can create completely opposite landscapes. It is possible to see an ever-changing landscape sequence in endless variations in which there are land, water, plants, animals and most importantly human communities on earth. Nature views are changing, so the landscape has a dynamic structure that is changed by natural, social, cultural and economic effects. Every element, living and inanimate, is part of the landscape. For this reason, the absence of each piece taken from nature is felt by us.

C Health Function of Green Areas

- One of the basic conditions for the city to be a healthy place is undoubtedly the size of the green areas in the city. In fact, it has long been known that green spaces have a positive effect on people's health. This fact was ignored for a long time in the face of the advantages of industrialization. However, green areas make a great contribution to the citizens in solving many health problems caused by urban life. The intense working pace of cities and the fact that individuals generally have to work indoors for a long time negatively affect the health of the employees. The green areas of the city offer opportunities for citizens to overcome these health problems. Frederick Law Olmsted, one of the architects of Central Park, considered one of the pioneers of landscape architecture, focused on the impact of green spaces on human health. Today, especially in developed countries, politicians and public health experts are trying to draw attention to the effect of green spaces on human health [17]. Studies show that green spaces have a calming effect on human behavior. On the other hand, the World Health Organization stated that green areas are a factor that reduces stress as well as providing clean air and quality physical activity to the society. Green areas contribute to softening the difficult and stressful qualities of urban life, albeit partially. It has been determined that especially the green areas of the cities contribute significantly to the individuals getting rid of stress, pain and tension. It has been determined that just a few minutes of people gazing at a landscape with trees, flowers and water can relieve tension, agitation and pain [18]. In any epidemic period, urban open and green areas protected from construction such as parks, sports fields, parking lots, squares, open areas of official, health education and religious facilities for multifunctional use should be of sufficient proximity and size to the population they will serve, and they should be in continuity and accessible. In landscape designs to be made in such areas, besides the necessity of making the road distances in sufficient dimensions, it is necessary to support it with stepped plants in the form of a protective wind curtain.
- From the past to the present, plants have been used gradually to cover visual pollution and to create wind curtains. By using the same function, especially on walkways, the distance between people can be maintained and clean environments can be created.

D The Importance of Landscape Design in Gardens

In Turkish gardens, trees have been used for functional purposes as well as for their spiritual purposes. Fruit trees and vegetable gardens were also used in these gardens for the needs of the household. Trees and various elements have been used to provide shade, and a hammock or swing has been installed on suitable trees to create a recreation and entertainment place for both children and adults.

Taşkan (2014), stated that as a result of the decrease in green areas today, residential gardens gain importance in order to satisfy people's desire for green, increase their relationship with nature, and at the same time add aesthetic value to houses. In the past and today, every poor or wealthy homeowner wanted to have a small or large courtyard or garden and stated that this desire was a major factor in the

thought of the garden as an integral part of the house and shaping it accordingly [19].

Any landscape planning or design in the gardens will beautify the environment and create spaces that will make people happy.

E The Importance of Green Space in the Pandemic Period

Green spaces have positive effects on the physical activity levels and health of people and children.

In order to ensure social distance, it is important that urban settlements are comfortable and calm and that the population density is distributed evenly. Creating comfortable and suitable social spaces for employees should be built in accordance with global standards of the amount of green space per individual. The importance of creating healthy living conditions in cities and designing green areas with appropriate plant species has increased in order to reduce chronic diseases.

Residential areas without gardens or green areas should not be planned. The shifting space of many sporting and cultural activities, which were not needed to be outdoors before, leads to the increased use of green areas. In this context, the protection and sustainability of the urban ecosystem also increases its importance. With the demands of the users such as spending more time in green areas, resting, lounging, exercising and doing sports activities, increasing the amount of green areas and ensuring the use of green areas in a way to protect social distance have come to the agenda. Nowadays, it is the amount of quality green areas that matters more than the amount of green space per person. The concept of quality green space, on the other hand, means that users can easily do many activities; that they can have a good time; will bring along areas that provide the sustainability of the ecosystem, appeal to all senses, and provide mental and physical satisfaction [20].

III Conclusion

Health and satisfaction increase in proportion to the naturalness of the landscape. The presence of green areas, the calmness of rural areas, bird sounds, flower and soil scent have a positive effect on human health which is especially important for patients. Accordingly, it is observed that sense or therapy gardens tend to develop near health-related institutions and organizations [21]. The design of the sensory garden, where the senses are stimulated with different colors, textures and forms of plants, is important in therapy gardens [22]. Therapy gardens must be implemented in accordance with both aesthetic and functional criteria. While most gardens satisfy people aesthetically, sensory and therapy gardens have been created with functionality to appeal to all five senses of human beings.

As a result of the environmental problems experienced today, many countries around the world are working on the definition, protection and management of their natural values. In this context, studies on the subject have increased with the European Landscape Convention, to which our country is also a party, but it is seen that it is not at sufficient level yet. For this reason, sanctions should be increased for more conscious use and protection of natural resources, and efforts should be focused on this issue.

With this study, it is revealed that cities should be planned and designed ecologically, whether there is a pandemic or not. Educational institutions, especially universities that provide environment, planning and design education, have important responsibilities in this regard. Meetings and training seminars should be held by university faculty members to raise awareness on the subject. In addition, local governments, non-governmental organizations and graduates who will be employed in the profession should be made aware of the importance and necessity of ecological approach in urban design.

References

[1] N Şahin, Ş 1999 Kırsal Peyzaj Planlaması Ankara Üniversitesi, Ziraat Fak Yayın N:1509 Ders Kitabı Ankara

[2] Bulut, Y , Atabeyoğlu, Ö 2010 Kent planlamasında peyzaj mimarlarının önemi III Ulusal Karadeniz Ormancılık Kongresi 20-22 Mayıs 2010 Cilt: IV Sayfa: 1494-1503

[3] Moore, E O (1981) A prison environment's effect on health care service demands Journal of Environmental Systems,11,17–34

[4] Ulrich, R S (1984) View through a window may influence recovery from surgery Science, 224: 420- 421

[5] Tennessen, C M & Cimprich, B (1995) Views to nature: Effects on attention Environmental Psychology, 15: 77-85

[6] Leather, P , Pyrgas, M , Beale, D & Lawrence, C (1998) Windows in the work place Environment and Behavior, 30(6), 739–762

[7] Kaplan, R (2001) The nature of the view from home Environment and Behavior, 33(4), 507–542 doi: 10 1177/00139160121973115

[8] Kuo, F E & Sullivan, W C (2001) Environment and crime in the inner city: does vegetation reduce crime? Environment and Behavior 33:343–367

[9] Cooper-Marcus, C & Barnes, M (1999) Healing Gardens: Therapeutic Benefits and Design Recommendations New York: John Wiley & Sons, 323-384

[10] Whitehouse, S , Vami, J W , Seid, M , Cooper-Marcus, C , Ensberg, M J , Jacobs, J R & Mehlenbeck, R S (2001) Evaluating a Children's Hospital Garden Environment Utilization and Consumer Satisfaction, Journal of Environmental Psychology 21, 301-314

[11] Pretty, J (2004) How nature contributes to mental and physical health Spirituality and Health Int 5(2):68 – 78

[12] Özdemir, A , & Fenkçi, M S (2016) İşitsel ve görsel peyzaj algısının hasta psikolojisindeki rolü Journal of Human Sciences, 13(2), 3022-3032 doi:10 14687/jhs v13i2 3793

[13] Richardson E A & Mitchell R (2010) Gender differences in relationships between urban green space and health in the United Kingdom Social Science Med , 71:568–575 doi: 10 1016/j socscimed 2010 04 015

[14] Tekeli, İ , 1980 "Türkiye'de Kent Planlamasının Tarihsel Kökleri, Türkiye'de İmar Planlaması", Mimarlık Fakültesi Yayınları, ODTÜ, Ankara

[15] Akdeniz, NS, Çelik, A , Ender, E , Zencirkıran, M 2015 Kentsel Ekosistemlerde Biyotopların Önemi: Bursa Kenti Örneği Tarım Bilimleri Araştırma Dergisi 8 (1): 40-43, 2015 ISSN: 1308-3945, E-ISSN: 1308-027X

[16] Benliay, A Gezer, NB 2019 Üniversite Yerleşkeleri İçin Çevresel Sürdürülebilirlik Dizinleri: Akdeniz Üniversitesi Örneği PEYZAJ - Eğitim, Bilim, Kültür ve Sanat Dergisi 2 (2019) 40-49

[17] Akpınar, A & Cankurt, M (2015) Türkiye'de kişi başına düşen yeşilalan miktarı ile ölüm oranı arasındaki ilişkinin incelenmesi Adnan Menderes Üniversitesi Ziraat Fakültesi Dergisi, 12(2), 101–107

[18] Karataş, A , Kılıç, S 2017 Sürdürülebilir Kentsel Gelişme ve Yeşil Alanlar SİYASAL: Journal of Political Sciences, 26(2): 53–78

[19] Bilir, ZI, 2019 Müstakil Konut Bahçesi Peyzaj Tasarımında Peyzaj Mimarı Ve Ev Sahiplerinin İletişimlerinin Sağlanması Ve Beklentilerinin Belirlenmesine Yönelik Bir Araştırma Ankara Üniversitesi Fen Bilimleri Enstitüsü Yüksek Lisans Tezi Ankara

[20] Seçkin, P 2020 Salgın Koşullarında Peyzaj Tasarımında Değişen Konfor Arayışları Spektrum Tasarım rehberi ISBN: 978-625-400-808-5 Sayı 01 06 2020

[21] Balode, L (2013) The design guidelines for therapeutic sensory gardens Research for Rural Development, 2, 114-119

[22] Erbino, C , Alessandro, T , Vagge, I & Ferrario, P S (2015) Guidelines fort he design of a healing garden for the rehabilitation of psychiatric patients Journal of Agricultural Engineering, XIVI (426) (pp 43-51)

Kaplan, R & Kaplan, S (1989) The experience of nature: a psychological perspective, pp 162–4 Cambrige University Press, New York

Kuo, F E (2001) Coping With Poverty: Impacts of Environment and Attention in the Inner City Environment and Behavior 33: 5–34

Taşkan, G (2014) Bartın kenti geleneksel konutlarındaki yapısal değişimlerin bahçe mekânı ve kullanılan bitki materyaline yansımaları, Yüksek Lisans Tezi, Bartın Üniversitesi Fen Bilimleri Enstitüsü

A Theoretical Study on Porous Functionally Graded Material Beams

E. MADENCİ

Necmettin Erbakan University, Konya/Turkey, emadenci@erbakan.edu.tr

***Abstract* – This study presents the static bending analysis of porous functionally graded material (FGM) beams. The Young's modulus and mass density of composite are assumed to be graded in the thickness direction according to two different distribution patterns. The classical beam theory is employed to derive the governing equations. The Ritz method is employed to obtain the transverse bending deflections. The results are compared with the beam made of porous material and reported in the literature.**

***Keywords* – Porosity, Functionally Graded Material, Beam, Variational Methods, Static Analysis**

I. Introduction

FUNCTIONALLY graded materials (FGMs), one of the new advanced composite materials first proposed by Japanese scientists, have since enjoyed great interest in research and engineering [1]. The FGMs are non-homogeneous composite materials that contain two or more components whose material composition changes continuously and smoothly along a specific direction(s) [2].

In contrast to plate and shell structures, limited studies investigate the mechanical properties of porous FGM beam. Wen [3] investigated on analytical solutions for deformation of a thick circular plate saturated by an incompressible fluid. Magnucki et al. [4] investigated bending and buckling of rectangular plate made of foam material. Magnucki and Stasiewicz [5] obtained an explicit expression for the critical load of FGM beam with graded porosity by using the principle of stationarity of the total potential energy.

There are important assumptions produced for the analysis of beams according to their geometric properties and the load effects applied accordingly. These; Euler-Bernoulli Beam Theory, Timoshenko Beam Theory and High-order Shear Deformation Beam Theories. The difference between these assumptions is due to the definition of the angle of shear. Ma and Wang [6] investigated nonlinear bending and post-buckling of circular functionally graded plates subjected to mechanical and thermal loadings based on the first-order shear deformation theories. Jabbari et al. [7] considered the stability of sandwich plate with piezoelectric layers and poroelastic core under uniform thermal and electrical field. They obtained their results based on the classical and first-order plate theories. Şimşek and Yurtcu [8] used the nonlocal Timoshenko and Euler–Bernoulli beam theory to examine the analytical solutions for the static bending and buckling of an FGM nanobeams. Grygorowicz et al. [9] studied the elastic buckling of three-layered beam consisting of metal foam core with varying mechanical properties, and applied a broken line hypothesis and a nonlinear hypothesis to the displacement field, respectively. Tomasz Belica et al. [10] investigated the dynamic stability of FG porous cylindrical shells subjected to combined loads. Magnuck-Blandzi [11] examined the non-linear dynamic stability and axi-symmetrical deflection and buckling of circular porous plates.

It should be noted that in the above-mentioned studies, only one specific porosity distribution was considered and no detailed discussion concerning the effects of different porosity parameters on the structural performance of porous beams was given. İn this study, the static bending of FGM porous beams with two different porosity distributions have been investigated. Theoretical formulations are within the framework of classical beam theory. Ritz method is employed to obtain the transverse bending deflection. The effects of porosity coefficient and slenderness ratio on the maximum deflection and associate are discussed.

II. Theoretical Formulation

A schematic of the FGM porous beam used in the present study is shown in Fig. 1. A FGM porous beam of thickness h and length L with two different porosity distributions along the thickness direction is shown in Fig. 1a for porosity distribution 1 and Fig. 1b for porosity distribution 2 [12].
Additionally, the continuous variation of the modulus of elasticity, shear modulus, and mass density due to the graded porosity are given by Eq. (1) and Eq. (2) as follows.

Model-1 (symmetric)

$$E(z) = E_1\left[1 - e_0 \cos(\frac{\pi z}{h})\right] \tag{1a}$$

$$G(z) = G_1\left[1 - e_0 \cos(\frac{\pi z}{h})\right] \tag{1b}$$

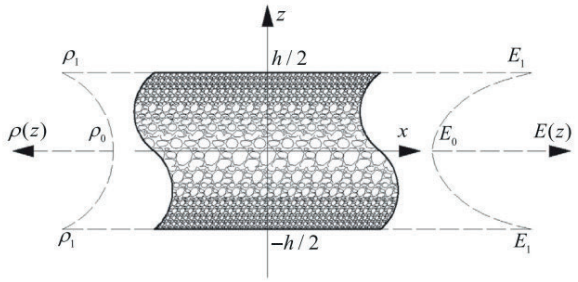
$$\rho(z) = \rho_1\left[1 - e_m \cos(\frac{\pi z}{h})\right] \tag{1c}$$

Model-2 (non-symmetric)

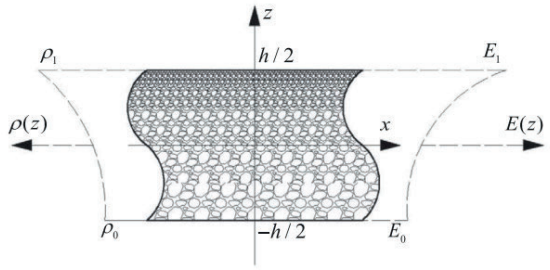
$$E(z) = E_1\left[1 - e_0 \cos(\frac{\pi z}{h} + \frac{\pi}{4})\right] \tag{2a}$$

$$G(z) = G_1\left[1 - e_0 \cos(\frac{\pi z}{h} + \frac{\pi}{4})\right] \tag{2b}$$

$$\rho(z) = \rho_1\left[1 - e_m \cos(\frac{\pi z}{h} + \frac{\pi}{4})\right] \tag{2c}$$



(a) Porosity distribution 1



(b) Porosity distribution 2

Fig. 1. Porosity distribution patterns [12]

where the porosity coefficient $e_0 = 1 - \frac{E_0}{E_1} = 1 - \frac{G_0}{G_1}$ ($0 < e_0 < 1$), E_0 and E_1 denote the minimum and maximum values, respectively, of the modulus of elasticity in the thickness direction of the beam. The porosity coefficient for mass density is defined as $e_m = 1 - \frac{\rho_0}{\rho_1}$ ($0 < e_m < 1$) in which ρ_0 and ρ_1 denote the minimum and maximum values, respectively, of mass density.

According to classical beam theory [13], the following displacement field is used which is neglected the effect of transverse shear strain

$$\varphi_i(x,y,z,t) = \sum_{j=0}^{N} (z)^j \varphi_i^j(x,y,t) \tag{3}$$

where "φ_i" is the ith component of displacement or stress, (x,y) the in-plane coordinates, z the thickness coordinate, t the time and "φ_i^j" are functions to be determined.

The closed form of the infinitesimal strain components is given by:

$$\varepsilon_{ij} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right) \tag{4}$$

when "φ_i" are displacements, then the equations governing "φ_i^j" are determined by the principle of virtual displacements or its dynamic version when time dependency is to be included

$$0 = \int_0^T (\delta U + \delta V - \delta K) \tag{5}$$

where δU, δV and δK denote the virtual work done by external applied forces and the virtual kinetic energy, respectively.

The linear constitutive model for infinitesimal deformation is referred to as the generalized Hooke's law. If the stress components are assumed to be linear functions of the components of strain, then the most general form of the linear constitutive equations for infinitesimal deformation is

$$\sigma_{ij} = C_{ijkl}\varepsilon_{kl} \tag{6}$$

where C is the 4th order tensor of material parameters and is termed stiffness tensor. We can obtain Euler equations as

$$\begin{aligned} &\delta u_0 \quad -\frac{dN}{dx} - f = 0, \quad 0 < x < L \\ &\delta w_0 \quad -\frac{d^2M}{dx^2} - q = 0, \quad 0 < x < L \end{aligned} \tag{7}$$

These governing equations are in the form of differential equations that are not always solvable by exact methods of solution. There is a number of approximate methods that can be used to solve differential equations such as "finite-difference methods, the finite element method, Ritz method etc". The Ritz method is employed to obtain the transverse bending deflection of the porous FGM beam with fixed-free (cantilever). The two displacement functions for beams with fixed-free boundary condition take the form of the following simple algebraic polynomials

$$\begin{aligned} u(\beta) &= \sum_{j=1}^{N} A_j \beta^j \\ w(\beta) &= \sum_{j=1}^{N} B_j \beta^j \end{aligned} \tag{8}$$

where N is the total number of polynomial terms, Aj and Bj are the unknown coefficients. The above two functions satisfy the geometric boundary conditions of the beams. Following the standard procedure of Ritz method, the minimizing the obtained total energy with respect to the unknown coefficients

$$\begin{aligned} \frac{\partial \dot{\Pi}}{\partial A_j} &= 0 \\ \frac{\partial \dot{\Pi}}{\partial B_j} &= 0 \end{aligned} \tag{9}$$

III. Numerical results

The material of the porous beam is assumed to be steel foam with E_1=200 GPa, v =1/3, q_1 = 7850 kg/m3. The cross section of beam is h =0.1 m, b=0.1 m (b is the width of the beam). It is assumed in this section that the uniformly distributed load is q = 1 x 10^4 N/m and the point load on the mid-span of the beam is F = 1 x 10^4 N, unless otherwise stated. The validation analysis is done through direct comparison between the present results and finite element results obtained by using commercial finite element software package ANSYS [12]. To simulate the graded distribution of elasticity moduli and mass density along the thickness direction, the beam is divided into a number of layers of same thickness.

Table 1: Dimensionless bending deflection under a distributed load (e_0=0.5).

L/h		Model 1	Model 2
10	Present	0.00083	0.0010
	ANSYS [12]	0.00083	0.00099
20	Present	0.01305	0.01576
	ANSYS [12]	0.01311	0.01572
50	Present	0.50988	0.61202
	ANSYS [12]	0.51007	0.61182

Table 1-2 show the dimensionless bending deflections $w = w_0/h$ under a distributed load q = 1 x 10^4 N/m, and a point load F = 1 x 10^4 N at the tip of the beam with varying slenderness ratio.

Table 2: Dimensionless bending deflection under a point load (e_0=0.5).

L/h		Model 1	Model 2
10	Present	0.00083	0.0010
	ANSYS [12]	0.00083	0.00099
20	Present	0.01305	0.01576
	ANSYS [12]	0.01311	0.01572
50	Present	0.50988	0.61202
	ANSYS [12]	0.51007	0.61182

IV. Conclusion

The static bending of FGM porous beams with cantilever boundary condition and two different porosity distributions have been investigated. Theoretical formulations are within the framework of classical beam theory. Ritz method is employed to obtain the transverse bending deflection. The effects of porosity coefficient and slenderness ratio on the maximum deflection is discussed.

References

[1] Rafiee M, Yang J, Kitipornchai S Large amplitude vibration of carbon nanotube reinforced functionally graded composite beams with piezoelectric layers Composite Structures 2013;96:716-25
[2] Chen D, Yang J, Kitipornchai S Free and forced vibrations of shear deformable functionally graded porous beams International Journal of Mechanical Sciences 2016;108-109:14-22
[3] Wen P The analytical solutions of incompressible saturated poroelastic circular Mindlin's plate Journal of Applied Mechanics 2012;79
[4] Magnucki K, Malinowski M, Kasprzak J Bending and buckling of a rectangular porous plate STEEL AND COMPOSITE STRUCTURES 2006;6:319-33
[5] Magnucki K, Stasiewicz P Elastic buckling of a porous beam Journal of Theoretical and Applied Mechanics 2004;42:859-68
[6] Ma L, Wang T Nonlinear bending and post-buckling of a functionally graded circular plate under mechanical and thermal loadings International Journal of Solids and Structures 2003;40:3311-30
[7] Jabbari M, Joubaneh EF, Mojahedin A Thermal buckling analysis of porous circular plate with piezoelectric actuators based on first order shear deformation theory International Journal of Mechanical Sciences 2014;83:57-64
[8] Şimşek M, Yurtcu H Analytical solutions for bending and buckling of functionally graded nanobeams based on the nonlocal Timoshenko beam theory Composite Structures 2013;97:378-86
[9] Grygorowicz M, Magnucki K, Malinowski M Elastic buckling of a sandwich beam with variable mechanical properties of the core Thin-Walled Structures 2015;87:127-32
[10] Belica T, Magnucki K Stability of a porous-cellular cylindrical shell subjected to combined loads Journal of Theoretical and Applied Mechanics 2013;51:927-36
[11] Magnucka-Blandzi E Non-linear analysis of dynamic stability of metal foam circular plate Journal of Theoretical and Applied Mechanics 2010;48:207-17
[12] Chen D, Yang J, Kitipornchai S Elastic buckling and static bending of shear deformable functionally graded porous beam Composite Structures 2015;133:54-61
[13] Reddy JN Mechanics of laminated composite plates and shells: theory and analysis: CRC press, 2004

Comparing The Precision of Global And Regional TEC Maps

S. BULBUL[1]

[1] Konya Technical University, Konya/Turkey, sbulbul@ktun.edu.tr

Layers surrounding earth sometimes have distorting effects on GNSS measurements. The ionosphere, is one of these layers, is the most important layer. The ionosphere layer has a structure that delays code measurements and speeds up phase measurements. Therefore, it is possible to determine the retarding and accelerating effects with GNSS measurements. GNSS measurements are used to model the effect of this layer. In the study, in 6 stations in CORS-Tr network, RIM-TEC maps were obtained using Global TEC maps. Later on, two-difference solution was realized by using both GIM-TEC and RIM-TEC maps. In the solutions, solar activity, geomagnetic storm values were taken into consideration. Based on the 173rd day of year 2015 in which there are three indexes, and the 174th day of year 2015 and the 7th day of year 2014, in where there are different indexes, the comparison of the coordinates calculated by taking these days and before and after 5 days has been made. As a result of the study carried out, it is seen that TEC maps produced from GIM TEC maps from Global TEC maps give exactly the same results with TEC maps produced regionally.

Keywords **– Bernese, GIM-TEC, RIM-TEC, precision, positioning**

I. INTRODUCTION

Earth's atmosphere consists of various gases and small particles. The atmosphere can be roughly defined as the region up to 1000 km above sea level. 99% of the atmospheric mass is below 30 km altitude [1], beyond 80 km altitude, the atmosphere contains ionized molecules and free electrons.

The ionosphere, which is located between 60 and 1100 km from the earth, is defined as the layer of atmosphere that surrounds the earth and is formed by gases ionized by solar rays [2-4]. Most of the ionosphere consists of neutral gases. Ionized gases are mostly formed as a result of ionization by short wave (ultraviolet and X-radiation) rays from the sun. The refractive index of the ionosphere, which changes according to the position, is the function of the carrier frequency for GNSS signals [4]. In general, the ionosphere can be shown changes according to geomagnetic location, geomagnetic activity, solar activity, number of sunspots, seasonality, local time, nuclear explosions, strong lightning and violent storms, ballistic missile flights and altitude [5-8].

The amount of ionization in the ionosphere layer is not similar in every region of the ionosphere layer. This situation can be explained by the fact that the energy carried by solar radiation is not the same in every region of the atmosphere. Ionization increases as it reaches the upper layers of the ionosphere. According to the ionosphere layer altitude; the layers can be listed as follows;

- Layer D (50-90 km),
- Layer E (90-140 km),
- Layer F_1 (140-240 km),
- Layer F_2 (240-500 km) (Şekil 1).

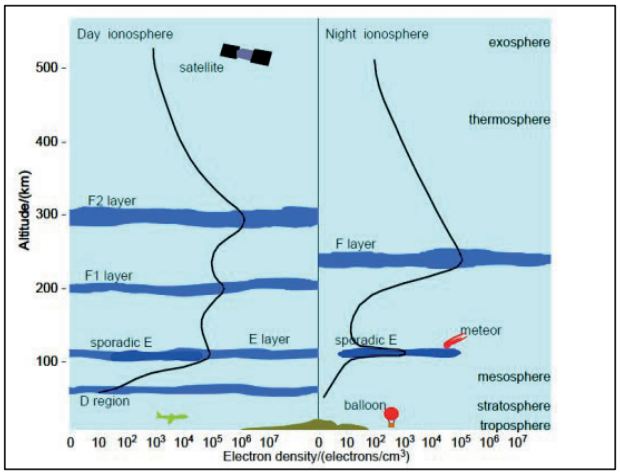


Figure 1. Sections of the ionosphere layer [9]

The ionosphere, or its parameter, TEC (Total Elecktron Content) on the signal pathway, differs depending on the number of free electrons. The number of free electrons depends on factors such as climatic conditions, time, geographic latitude, sun and geomagnetic change [10]. These changes in the ionosphere layer can be examined in two main sections as regular and irregular gradients.

Daily maximum and minimum changes in TEC can be given as an example to regular gradients. After all, there are gradients that depend on local time. Another example of regular gradients is the ionosphere changes occurring in the north-south and east-west directions [11]. Irregular gradients are traveling ionospheric disturbing effects [12].

The informations related to the ionosphere can be accessed in detail with the help of numerous GNSS observation stations which are located around the earth and GNSS satellites. Since the ionosphere is a scattering environment, TEC values can be determined with the help of L_1 and L_2 carrier signals sent from GNSS satellites. TEC values contain information about the global or regional ionosphere structure [13-14]

The Local (regional) TEC map is obtained by applying Taylor expansion to the linear combination of L4, which is the difference of the phase measurements (L1, L2).

$$L_4=L_1-L_2 \quad (1)$$

The regional Taylor expansion has been fall short of modeling the global ionosphere effects. Spherical harmonic expansion is used for this [15-16].

The point where the GNSS signal between the station and the satellite cuts this infinitely thin layer is called the puncture point of the ionosphere. The TEC value expressing the satellite-receiver path at the ionosphere puncture point is called STEC (TEC in the slope direction), and the vertical projection of this value is VTEC (TEC in the vertical direction). STEC values are converted into VTEC (Vertical Total Electron Content) values that define the electron activity in the vertical direction in order to determine the total electron density in the ionosphere (Figure 2).

Conversion between STEC and VTEC values is achieved by the reduction function.

$$\sin(z') = \frac{R_E}{R_E+H}\sin(z) \quad (2)$$

$$MF_1 = \frac{STEC}{VTEC} = \frac{1}{\cos(z')} \quad (3)$$

In equations (2) - (3), MF: Mapping Function, R_E denotes the radius of the earth and H denotes the height from the earth face.

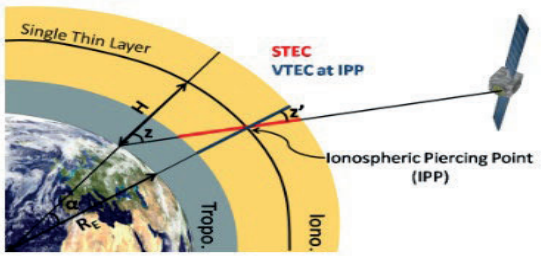


Figure 2. STEC and VTEC [17]

II. Material and Method

Electron density and TEC changes of the ionosphere are investigated with the help of ionospheric models. Ionospheric models are divided into two as theoretical and experimental. Experimental models are models made with the help of a computer program. The most commonly used software in making these models are IRI, Bernese, Gamit-Globk and JPL.

In June 1998, the International GNSS service (IGS) started an international project to calculate daily different global ionosphere maps (GIMs) using GPS data. The institutions that produce the global ionosphere TEC map in the world are named the Ionosphere Associate Analysis Center (IAAC) and are listed below [18-21]:

- CODE: The Centre for Orbit Determination in Europe, Bern, Switzerland,
- DLR : Deutsches Zentrum für Luft- und Raumfahrt, Germay,
- ESA: The European Space Agency, Darmstadt, Germany,
- JPL : Jet Propulsion Laboratory, Pasadena, CA, U.S.A.,
- NOAA : National Oceanic and Atmospheric Administration, Silver Spring, U.S.A.,
- NRCan : Natural Resources Canada, Ottawa, Ontario, Canada,
- ROB : Royal Observatory of Belgium, Brussels, Belgium,
- UNB : University of New Brunswick, Fredericton, N.B., Canada,
- UPC: Universitat Politècnica de Catalunya, Barcelona, Spain
- WUT : Warsaw University of Technology, Warsaw, Poland.

Global ionosphere map (GIM) is published in IONEX (IONosphere map EXchange) format. Global ionosphere maps in IONEX format are produced at 2-hour intervals in UTC time. For TEC values, the increase in longitude is 5° and the increase in latitude is 2.5° [11], [19]. The accuracy of TEC values broadcast from IONEX format varies between 2-8 TECU.

A. Centre for Orbit Determination in Europe (CODE)

CODE Global Ionosphere Maps (GIM) produces TEC values on a daily basis at 2-hour intervals using data from about 200 GPS stations of IGS and other institutions [22-23]. Vertical Total Electron Content (VTEC) up to 15 degrees was modeled using spherical harmonics expansion in the solar-geomagnetic reference frame. For the calculation of TEC values, the earth is considered to have a radius of 6371 km [21]. The Global Ionosphere model in IONEX format produced in the CODE analysis center is given in Figure 3.

```
     1.0            IONOSPHERE MAPS     GNSS                IONEX VERSION / TYPE
ADDNEQ2 V5.1        AIUB                18-AUG-12 20:47     PGM / RUN BY / DATE
CODE'S GLOBAL IONOSPHERE MAPS FOR DAY 227, 2012             COMMENT
Web page:        www.aiub.unibe.ch/content/ionosphere/      DESCRIPTION
Data archive:    ftp.unibe.ch/aiub/CODE/                    DESCRIPTION
                 www.aiub.unibe.ch/download/CODE/           DESCRIPTION
  2012     8    14     0     0     0                       EPOCH OF FIRST MAP
  2012     8    15     0     0     0                       EPOCH OF LAST MAP
  7200                                                      INTERVAL
    13                                                      # OF MAPS IN FILE
  NONE                                                      MAPPING FUNCTION
    10.0                                                    ELEVATION CUTOFF
One-way carrier phase leveled to code                       OBSERVABLES USED
   274                                                      # OF STATIONS
    56                                                      # OF SATELLITES
  6371.0                                                    BASE RADIUS
     2                                                      MAP DIMENSION
   450.0 450.0   0.0                                        HGT1 / HGT2 / DHGT
    87.5 -87.5  -2.5                                        LAT1 / LAT2 / DLAT
  -180.0 180.0   5.0                                        LON1 / LON2 / DLON
    -1                                                      EXPONENT
TEC/RMS values in 0.1 TECU; 9999, if no value available     COMMENT
Peak TEC values for the included maps:                      COMMENT
 589  584  643  639  680  625  541  548  624  660  735  769 COMMENT
 704                                                        COMMENT
                                                            END OF HEADER
     1                                                      START OF TEC MAP
  2012     8    14     0     0     0                       EPOCH OF CURRENT MAP
    87.5-180.0 180.0   5.0 450.0                           LAT/LON1/LON2/DLON/H
 117  117  118  118  118  119  119  119  119  119  119  119  119  118  118
 117  117  116  116  115  114  114  113  112  112  111  111  110  110  109  109
 109  109  108  108  108  108  108  108  108  108  109  109  109  109  109  109
 110  110  110  110  110  110  110  111  111  111  111  111  112  112  112  112
```

Figure 3.CODE Global Ionosphere Model

III. APPLICATION

In the study carried out, the effect of TEC values on point coordinates was investigated by using TEC values obtained from global maps and regionally produced TEC maps. While investigating the effect of GIM and RIM TEC maps, the solutions realized depending on the situation of geomagnetic effects (Kp, Dst and F10.7 indexes) were taken into consideration. For this purpose, AKSR, BOGZ, KAYS, KIRS, NEVS and NIGD stations of CORS-Tr network, were used. Informations about the stations are given in Table 1.

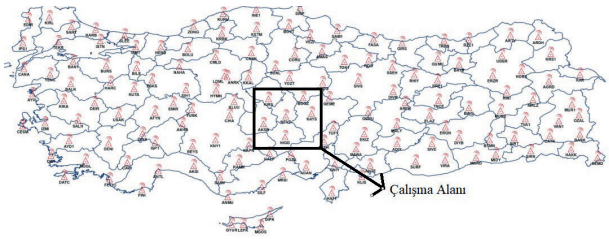


Figure 4. CORS-Tr stations used

Table 1. Information on the stations used

PI	latitude (o)	Longitude(o)	Ellipsoidal height (m)	Receiver	Antenna
AKSR	38.37139886	33.99845065	963.1069	TRIMBLE NETR5	TRM55971.00
BOGZ	39.19459561	35.25537335	1066.2444		
KAYS	38.70936915	35.52472513	1097.5044		
KIRS	39.16561511	34.15578928	1022.5379		
NEVS	38.61756646	34.70323048	1250.6028		
NIGD	37.95978564	34.67971884	1209.2125		

Among the selected days, it was determined that the F10.7, dst and Kp indices were high on the 173rd day of 2015, the F.10.7 index was low on the 174th day of 2015, and only F10.7 index was high on the 7th day of 2014, while the other indices were low. The graphic display of the selected days is shown in Figure 5-7.

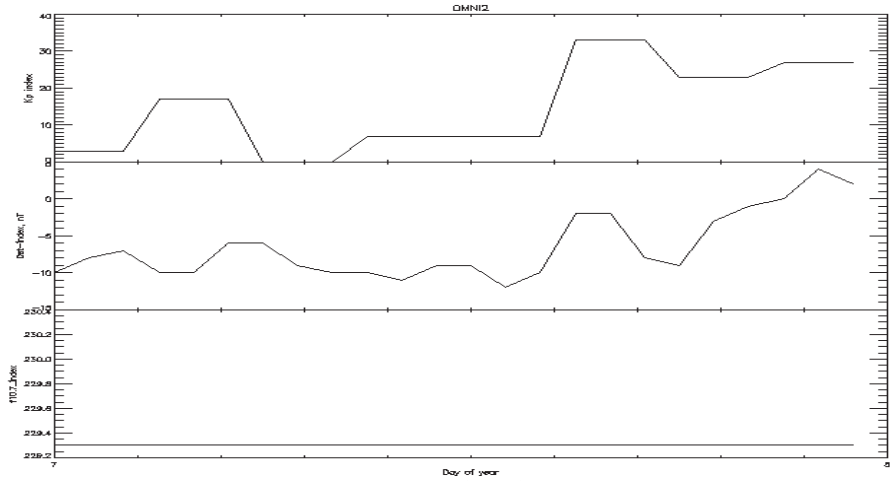


Figure 5. Indices for the 7 days of 2014

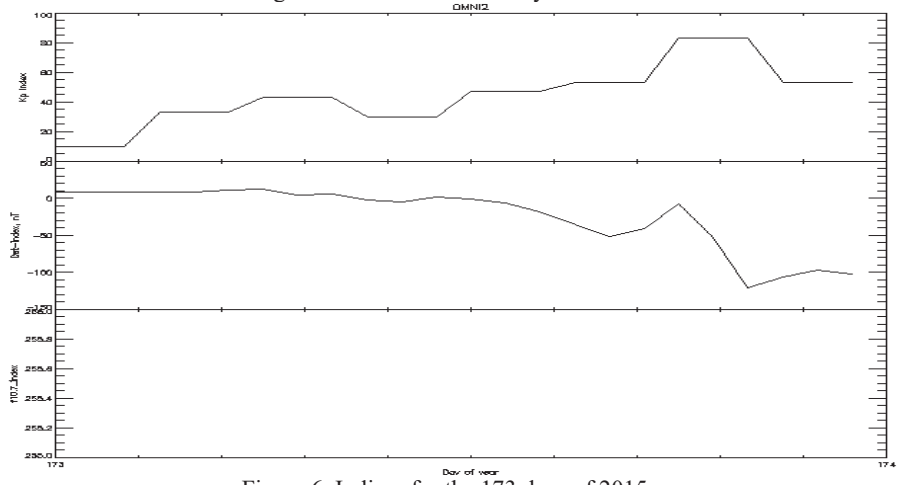


Figure 6. Indices for the 173 days of 2015

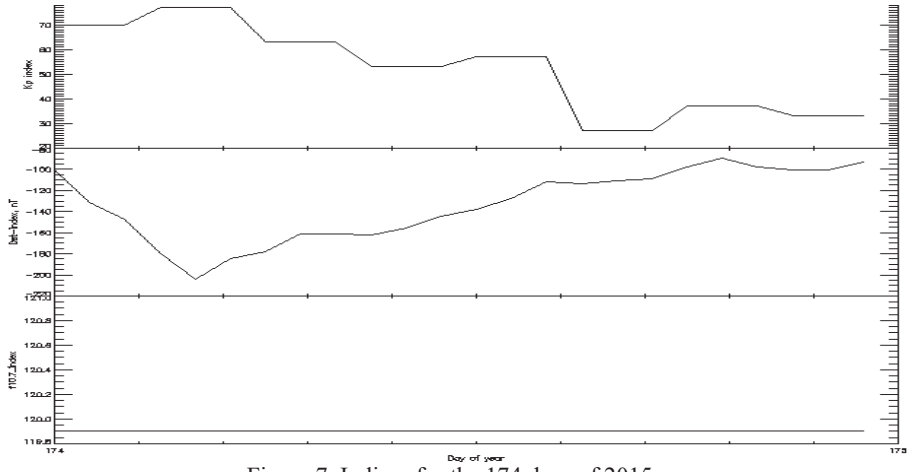


Figure 7. Indices for the 174 days of 2015

A. A. Production of Regional TEC Maps

PPP_DEMO.PCF automatic solution strategy of Bernese v5.2 Scientific GNSS software was used to obtain regional TEC values. In this context, the solution strategy for Bernese v5.2 Scientific GNSS software is shown in Figure 8.

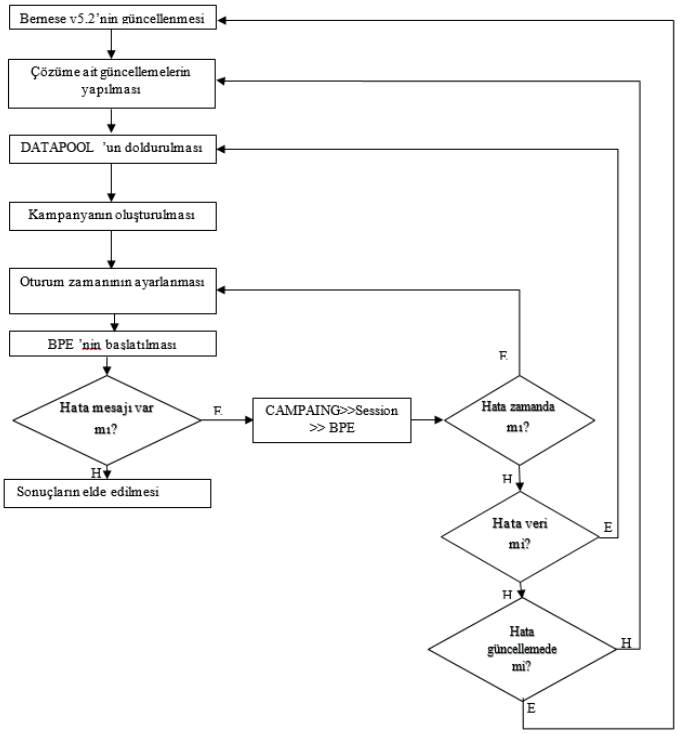


Figure 8.PP_DEMO.PCF automatic solution module workflow

Changes in TEC values calculated by taking into account the geomagnetic effects by using Global and Regional TEC maps are shown in Figure 9-14..

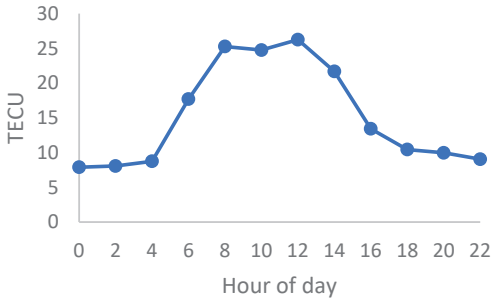


Figure 9.TEC values of Global TEC maps for the 7th day of 2014

As seen in Figure 9, TEC values reach the minimum level at 00:00 and the maximum level at 12:00. In addition, it reaches a minimum of 7.9 TECU and a maximum of 26.3 TECU during the day.

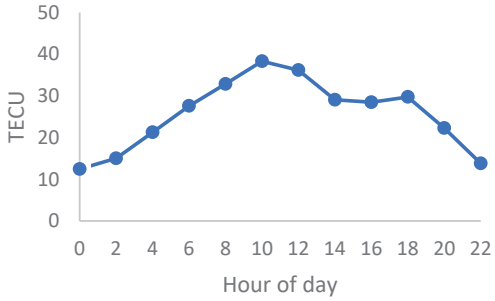


Figure 10.TEC values of Global TEC maps for the 173th day of 2015

In Figure 10, it moves similarly to the previous day, reaching a minimum of 12.5 TECU and a maximum of 38.3 TECU during the day.

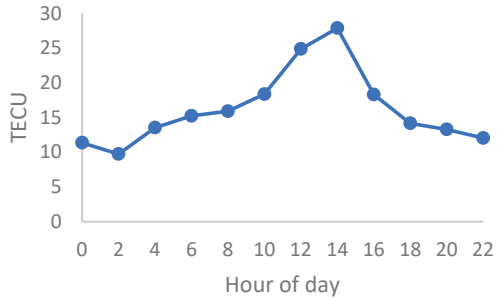


Figure 11.TEC values of Global TEC maps for the 174th day of 2015

In Figure 11, it moves similarly to the previous day, reaching a minimum of 9.7 TECU and a maximum of 27.9 TECU during the day. When Figure 9-11 is examined together, it is seen that the TEC values reached the maximum levels on the 173rd day of 2015 when geomagnetic effects were high.

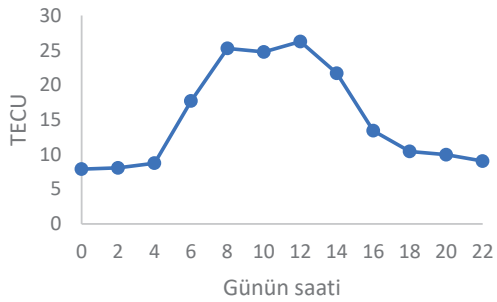


Figure 12.TEC values of RIM TEC maps for the 7th day of 2014

As seen in Figure 12, TEC values reach the minimum level at 00:00 and the maximum level at 12:00. In addition, it reaches a minimum of 7.9 TECU and a maximum of 26.2 TECU during the day.

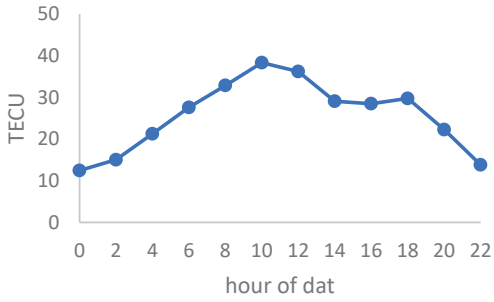


Figure 13.TEC values of RIM TEC maps for the 173th day of 2015

In Figure 13, it moves similarly to the previous day, reaching a minimum of 12.5 TECU and a maximum of 38.3 TECU during the day.

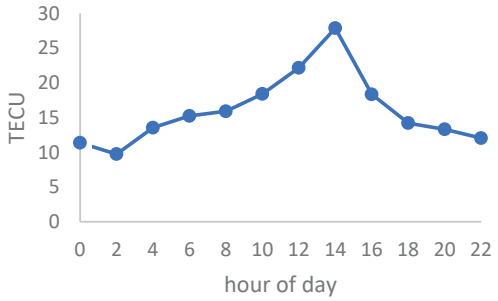


Figure 14.TEC values of RIM TEC maps for the 174th day of 2015

In Figure 14, it moves similarly to the previous day, reaching a minimum of 9.7 TECU and a maximum of 27.8 TECU during the day. When Figures 11 and 14 is examined together, it is seen that the TEC values reached the maximum levels on the 173rd day of 2015 when solar activity, magnetic storm and solar activity were high.

After calculating the regional TEC values for each day, the Regional TEC values for each day were obtained by adding solutions 5 days before and 5 days after the days of geomagnetic effects. The IONEX file of each day was then converted to the Bernese format for use in Double–Difference Network Processing. The Double–Difference Network Processing for each day was realized with the automatic solution module RNX2SNX.PCF of Bernese v5.2 Scientific GNSS software. By using the RNX2SNX.PCF automatic solution module, the solutions for each day were analyzed with both TEC values obtained from Global maps and TEC maps obtained regionally.

To determine the point location precisions,

$$\begin{aligned} \varepsilon_X &= X_{ort} - X_i \\ \varepsilon_Y &= Y_{ort} - Y_i \\ \varepsilon_Z &= Z_{ort} - Z_i \end{aligned} \tag{4}$$

With Equation (4), the difference of the coordinates from the mean value is calculated. Then, with the help of the calculated differences, the root mean square errors (rmse) in each axis direction was calculated with equation (5).

$$\begin{aligned} m_X &= \pm\sqrt{\frac{[\varepsilon_X \varepsilon_X]}{n}} \\ m_Y &= \pm\sqrt{\frac{[\varepsilon_Y \varepsilon_Y]}{n}} \\ m_Z &= \pm\sqrt{\frac{[\varepsilon_Z \varepsilon_Z]}{n}} \end{aligned} \tag{5}$$

In Equation (5); n: number of measurements (1,2,3,…, 23), $X_{ort}, Y_{ort}, Z_{ort}$; average coordinate of each station, X_i, Y_i, Z_i; show the coordinates of each day.

Difference statistics calculated with the help of both Global TEC maps and Regional TEC maps of each station are shown in Table 2.

Table 2.Difference statistics obtained from Global and Regional TEC maps

	Global maps			Regional Maps		
ID	ε_X	ε_Y	ε_Z	ε_X	ε_Y	ε_Z
	1.24	0.57	1.18	1.24	0.57	1.18
AKSR	-2.47	-1.00	-1.11	-2.47	-1.00	-1.11
	1.83	0.68	1.17	2.33	1.00	1.37
BOGZ	-3.75	-2.51	-2.43	-2.55	-0.88	-2.23
	1.04	0.74	1.33	1.04	0.74	1.33
KAYS	-1.52	-0.50	-1.36	-1.52	-0.50	-1.36
	0.75	0.71	1.14	0.75	0.71	1.14
KIRS	-1.85	-0.73	-1.03	-1.85	-0.73	-1.03
	0.91	0.79	1.28	0.91	0.75	1.16
NEVS	-1.63	-0.62	-1.16	-1.64	-0.66	-1.28
	1.19	1.15	1.43	1.19	1.15	1.43
NIGD	-1.89	-0.64	-1.31	-1.89	-0.64	-1.31

Table 3 shows the rmse in each coordinate axis by making use of difference statistics belong to both Global TEC maps and Regional TEC maps.

Table 3. The rmse obtained from Global and Regional TEC maps

	Global maps			Regional maps		
ID	m_X	m_Y	m_Z	m_X	m_Y	m_Z
AKSR	1.13	0.35	0.60	1.13	0.35	0.60
BOGZ	1.68	0.73	0.79	1.63	0.65	0.78
KAYS	0.92	0.29	0.79	0.92	0.29	0.79
KIRS	0.75	0.30	0.50	0.75	0.30	0.50
NEVS	0.87	0.32	0.69	0.87	0.32	0.70
NIGD	1.06	0.41	0.69	1.06	0.41	0.69

IV. Conclusion

In the study carried out, it is aimed to investigate the effect of Global and Regional TEC maps on location accuracy. In this context, RINEX data of 6 stations were selected from CORS-TR stations. The effects of Kp, F10.7 and Dst indices have been also investigated in determining the days. In this context, the 173rd day of year 2015, in which there were all three indexes, and the 174th day of year 2015 and the 7th day of year 2014, in which there were different indices were incloued to processing. Then, the effect of Regional TEC maps produced by including all three days in the analysis 5 days before and 5 days after, on point location accuracy was investigated.

When Figure 6-8 showing the TEC values obtained from the Global TEC maps and Figure 9-11 showing the TEC values generated from the Regional TEC maps are examined together, it is seen that both Global and Regional TEC maps show similar movements for the selected region.

After the regional TEC maps are obtained, it is seen that coordinates,obtained via the Double–Difference Network Processing, are similar to the coordinates obtained by using both Global TEC maps. When Table 2 has been examined, it is seen that the difference statistics are similar.

After the difference statistics are determined, when Table 3, which shows the rmse calculated based on the difference statistics, is examined, it is seen that the rmse calculated in each coordinate axis direction are almost exactly the same.

As a result of the study carried out, it is seen that TEC maps produced from GIM TEC maps give exactly the same results with the TEC maps produced regionally in selected region.

References

[1] Memarzadeh, Y , 2009, Ionospheric modeling for precise GNSS applications, Ph D Thesis, Delft University of Technology, Netherlands

[2] Arıkan, F , Erol, C ve Arıkan, O , 2003, Regularized estimation of vertical total electron content from Global Positioning System data, Journal of Geophysical Research: Space Physics, 108 (A12), 1469-1480

[3] Turel, N , Aktas, E ve Arikan, F , 2007, TEC Statistics and Correlogram Applications, Signal Processing and Communications Applications, 2007 SIU 2007 IEEE 15th, 1-4

[4] Aysezen, M Ş , 2008, Türkiye İçin IONOLAB-TEC Kullanılarak GPS Tabanlı TEİ Ve Alıcı Yanlılığı Veri Tabanı Hazırlanması, Yüksek Lisans Tezi, Zonguldak Karaelmas Üniversitesi Fen Bilimleri Enstitüsü, Zonguldak

[5] Wild, U , 1994, Ionosphere and geodetic satellite systems: permanent GPS tracking data for modelling and monitoring, Geod -Geophys Arb Schweiz, Vol 48, 48

[6] Hawarey, M ve Ayan, T , 2004, Uzay mekiği tırmanışı, deprem, ve füze fırlatılışından kaynaklanan TEC değişimlerinin GPS ile belirlenmesi, İTÜ Dergisi, 3 (2-3-4-5)

[7] Namgaladze, A A , Zolotov, O V , Karpov, M I ve Romanovskaya, Y V , 2012, Manifestations of the earthquake preparations in the ionosphere total electron content variations, Natural Science, 4 (11), 848-855

[8] Liu, M , Luo, G ve Wang, H , 2014, The 2013 Lushan earthquake in China tests hazard assessments, Seismological Research Letters, 85 (1), 40-43

[9] URL-1,2020, Web adresi URL: http://www idc-online com/technical_references/pdfs/electronic_engineering/Day_and_night_structure_of_ionosphere pdf (Ziyaret Tarihi 18 Mart 2020)

[10] Parkinson, B W ve Spilker, J J , 1996, Global Positioning System: Theory and Applications, p

[11] Arslan, N , 2004, GPS ile İyonosfer Toplam Elektron Yoğunluğu Değişimlerinin Koordinatlara Etkilerinin Araştırılması, Doktora Tezi, Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü, İstanbul

[12] Warnant, R , 2002, Atmospheric Perturbations on GNSS Signals and Their Influence on Time Transfer, XXVIIth General Assembly of the International Union of Radio Science, Maastricht, The Netherlands

[13] Davies, K ve Hartmann, G , 1997, Studying The Ionosphere With The Global Positioning System, Radio Science, 32 (4), 1695-1703

[14] Fedrizzi, M , Langley, R B , Komjathy, A , Santos, M C , de Paula, E R ve Kantor, I J , 2001, The Low-Latitude Ionosphere: Monitoring Its Behavior With GPS, Proceedings of ION GPS, Salt Lake City, Utah, USA

[15] Schaer, S , Beutler, G , Mervart, L , Rothacher, M ve Wild, U , 1995, Global and regional ionosphere models using the GPS double difference phase observable, IGS Workshop on Special Topics and New Directions, Germany

[16] Ateş, H B , 2011, TUSAGA-AKTİF GPS Ağ Verileri İle Bölgesel İyonosferik Modelin Oluşturulması, Yüksek Lisans Tezi, Gebze Yüksek Teknoloji Enstitüsü Mühendislik Ve Fen Bilimleri Enstitüsü, Gebze

[17] URL-2,2020, Web adresi URL : http://gnss be/ionosphere_tutorial php (Ziyaret Tarihi 4 Nisan 2020)

[18] Feltens, J ve Schaer, S , 1998, IGS Products for the Ionosphere, Proceedings of the 1998 IGS Analysis Center Workshop Darmstadt, Germany

[19] Orús Pérez, R , 2005, Contributions on the improvement, assessment and application of the Global Ionospheric VTEC Maps computed with GPS data, Ph D dissertation, Universitat Politècnica de Catalunya, Barcelona, Spain

[20] Orús Pérez, R , 2005, Contributions on the improvement, assessment and application of the Global Ionospheric VTEC Maps computed with GPS data, Ph D dissertation, Universitat Politècnica de Catalunya, Barcelona, Spain
[21] Başçiftçi, F , İnal, C , Yıldırım, Ö ve Bülbül, S , 2017, Determination of Regional TEC Values by GNSS Measurements, A Case Study: Central Anatolia Sample, Turkey Surveying the world of tomorrow – From digitalisation to augmented reality, FIG Working Week 2017 Helsinki/Finland
[22] Liu, Z , Skone, S , Gao, Y ve Komjathy, A , 2005, Ionospheric modeling using GPS data, Gps Solutions, 9 (1), 63-66
[23] Todorova, S , Hobiger, T , Weber, R ve Schuh, H , 2003, Regional Ionosphere Modelling With GPS and Comparison With other Techniques, proceeding of: Proceedings of the Symposium" Modern Technologies, Education and Professional Practice in the Globalizing World", November, 06-07

Evaluation of the Relationship Between Road and Landscape in terms of Landscape Planning Criteria

G.SANDAL ERZURUMLU and M.F. TEKİNALP.

Niğde Ömer Halisdemir University, Niğde/Turkey, gpeyzaj@gmail.com
Düzce University, Düzce/Turkey, mertkantekinalp@gmail.com

Abstract. **Adverse effects of rapidly developing technology and environmental pressures created by improper use of land on natural resources are among the important problems in the world agenda. Roads are of great importance in terms of touristic, cultural and socio-economic structure in developing countries. Roads should be in harmony with the modern traffic and in an order that can serve the needs without disturbing the natural structure of the area they pass through. Road planning shoud be made during the route selection phase, not after the construction phase is completed and opened to traffic. During the road construction phase, the project speed of the road, the route of the road, parallelism, excavation and fillings and planting should be complied with landscape planning principles. In this study, the aforementioned criteria have been considered, and suggestions have been made for the relationship between road and landscape.**

Keywords: **Landscape, Road Landscape, Highways planning, Planting, Green road**

I.INTRODUCTION

Situated in the category of developing countries, Turkey has a high rate of population growth and urbanization. Especially in the last 20 years, almost doubled number of urban population in Turkey has led to the formation of dense residential areas. Cities expanded in the north-south and east-west direction, and settlements spread over large areas. In this process, the increase in living standards, the expansion of industrial zones, the increase in the importance people give to education and cultural-artistic activities have led to an increase in transportation needs [1].

Land transport has a 95% share in freight and passenger transport all over the world. It is constantly widespread and has become one of the most important construction investments. The increase in the world population, the development of the industry and the increase in the need for raw materials increase the demand for roads. The roads can also be defined as all of the structures built according to geometric standards in order to bring the natural ground together and ensure that motor vehicles move in safety, comfort and desired speed conditions [2,3,4] Roads consists of traffic, ecology, recreation, culture, education, etc. functions.

Roads are the places where the land use is the most intense and the human impact is the highest. With the development of the highways industry, increasing migration from rural to urban to work in production centers, housing, transportation, commercial uses, fossil fuel consumption, deforestation, wrong land use, pollution caused by greenhouse gases released into the atmosphere, production and consumption pressures, these areas becomes the areas where environmental problems are concentrated. As a result of this, the decreasing green areas in urban spaces break human-nature relations, which leads to the emergence of increasing ecological problems in cities. These problems show that in today's cities, urban development is left to an uncontrolled development and change, and it lacks an ecological basis. In order to help reduce these harmful effects, it has been revealed that transportation areas should be planned with the understanding of providing environmentally sensitive urban development with vegetation on highways and roads. According to Korku et al. (2017), green infrastructure, green road, green wedge etc. systems are some of the studies done in this direction.

Radial roads suitable for the Baroque style, with their five-row trees, have passed from the cities and reached the countryside. The roads with regular lines on both sides, called "Allée", started with this period and the aforementioned style [7,8]. In 19th century Europe, the road trees aimed to beautify the landscape and prevent dust. Fruit trees were used for a long time for this purpose [7,8,9]. Erosion control started on highways in 1920-1930 [7,9,10]. In the world, the first examples of highway landscape planning studies started in the 1930s. Landscape architects and forest engineers examined the possibilities of using vine logs and shrubs in order to prevent erosion that occurred on highways during these years [9,10]. The first study on highway landscape planning in our country is Saatçioğlu's (1960) study, which examines the silvicultural and aesthetic features of roadside trees. Later, Akdoğan (1967) determined the features and principles of Ankara-Istanbul highway landscape planning [10]. Tanrıverdi (1975) also has a study on highway afforestation [14].

II HIGHWAYS PLANNING AND DESIGN PRINCIPLES

Project/Design speed of the road

The road is a strip of land open to public use for all kinds of vehicles and pedestrians. It is desirable that a road be safe, fast, comfortable and convenient throughout its service life. A road with these features is realized by applying fundamental engineering criteria in planning, design, construction and maintenance - repair - renewal stages. The most important factor in choosing the geometric standards of a road is the

project speed and the road class. The road class is determined according to the amount of traffic that will use the road, the type of traffic and the topographic characteristics of the land. In addition, the condition of the ground, climatic conditions, land use pattern and above all, economic possibilities are effective in choosing road standards [15]. Standards such as project speed, lane width, shoulder width, platform width are determined according to these service levels.

In the project specifications of the Ministry of Agriculture and Forestry, it is stated that the lane width in rural roads should be between 2.5 - 3.5 meters depending on the class of the road and the project speed. The roads will be designed to serve the parcels both sided in some places or unilaterally in others. According to the standards of our project determined by considering the average daily traffic amount, road class and project speed; The roads that will be parallel to the evacuation channels will have a platform width of 6 meters, blind roads 4 meters, and all other roads 7 meters. In particular, in order to prevent traffic accidents, loss of life and property and to ensure road safety, these criteria will be taken into account when planning in-field roads in the project area. [16].

A. The Route of the Road

The purpose of speedways is to create a fast and secure link between countries, industry and trade centers or cities. However, experiments have shown that the safest distance between two points is not straight lines. A road with gentle curves helps to keep the driver awake, enjoy the view and provide good visibility in both directions. In their study by Koç and Şahin (1999), the conditions deemed necessary for the route are listed below as:

- The horizontal curve in a curved section of the road and the vertical curve should be in harmony in length,
- Combination of short curves and lines should be avoided, curves in the same direction should not be connected with lines
- Additional slopes of more than 1% on the general slope make the view ugly
- Correct road sections should be gradually, transitioning between curves

It is appropriate to give the route (S) curves in relation to the vertical curves.

B Parallelism

Parallelism is a view that emerges from the parallel formation of the road, refuge and expropriation border lines. Parallelism is most effective in applications of roads with high geometric standards in rough terrain. Especially in contradiction with the size of the local landscape, when a highway is passed through a miniature scale nature, the road becomes an obvious and ugly trail due to parallelism. Parallelism can be avoided by expanding the expropriation boundaries from time to time on the highways, by changing the width of the central refuge 12, by making changes between the levels of the round-trip road or by considering the departure and arrival routes separately, thus establishing a more ideal connection between nature and the road [18,19].

C Excavation and filling

Very steep cuts and slopes always present a risk of sliding in rainy months. Retaining walls are generally made of concrete to prevent sliding. However, walls often cannot prevent erosion completely [10]. The vegetation in the excavation and filling slopes formed with very steep slopes makes it difficult to hold the top soil in terms of erosion control due to the steepness, and the deviations from planting techniques during planting and the disadvantages that make adequate maintenance difficult. Excavation and fill slopes should be created with a slope of at least 1/3, if possible, with a slope of 1/4 [20].

D. Preparation of Highway Landscape Plans

It is necessary to prepare a plantation plan in order to determine the locations of trees and shrubs in afforestation of the highway route, to make cost calculations and to make a correct application. The highway must comply with the character of the landscape in which the landscape facilities are located, the location of the landscape, the road type and traffic rules.

- Duties and functions of road trees and bushes:
- Plant materials to be used in road afforestation are examined in 4 different groups.

a. Duties in terms of construction technique:
Stabilization of the soil
Preventing landslides
Snow fence
Trench against avalanche, rock and stone falling
b. Duties in terms of traffic technique:
Good visibility of the road
c. Cover against headlights
Preventing or mitigating accidents and mistakes
Windbreak
Hiding objects that want to be left out of sight
d. Duties in the Landscape:
Creating a green environment again
Connecting different landscape units on the route
Merge, alter and create a new landscape area
e. Biological functions
Conserving wildlife
Providing flora for bees
Small microclimate formation

E. *Considerations in Highways Planning*

Inclusion of existing trees and shrubs in the plan

Previously developed trees and shrubs in the landscape should be used as much as possible. It should be ensured that the existing plants during road construction are not damaged.

Passing road routes through the forest area:
The road passing by contacting the forest
Crossing the road by partially cutting the forest
Passing the road directly through the forest
Road elevation
Free placement of groups
Signal-effective tree planting
Central refuge

Special cases for roads passing through forest areas:
Trees falling on the road with the effect of wind
Burn due to sun,
Frozen road

F Planting

The afforestation to be made on the road route is a complex issue that concerns the landscape through which the road passes in one respect, and the road itself, with transportation safety and driving pleasure [17].

Functions of road trees: Between the years 1890-1900, when the automobile began to be seen on the highways in the USA, forestation was carried out to prevent the dust of the roads and to provide shade in hot regions. The functions of vegetation in terms of traffic technique in the transport system on highways are listed as follows:

- *Road afforestation and functions*
- *Avoiding headlights*: Generally in the middle refuges of the roads, if the road passes through a flat area, a 2.5m high refuge afforestation will provide protection. On the road that passes through the valley pit, a higher mass of green area is required (Figure 1).

Figure 1.Hedge Plants in road [21].

Preventing or mitigating accidents

It is ensured that the harmful effects of the headlight lights of the vehicles coming from the opposite lane are prevented by planting the bends in divided roads. Similar methods are used to prevent the residents living around the road from being disturbed. Existing bends can be highlighted with vegetation from the hill where the slope ends, to ensure the road understanding by the driver. In addition, it is possible to protect the driver and passengers from the irritating effect of sunlight with vegetation (Figure 2) [22].

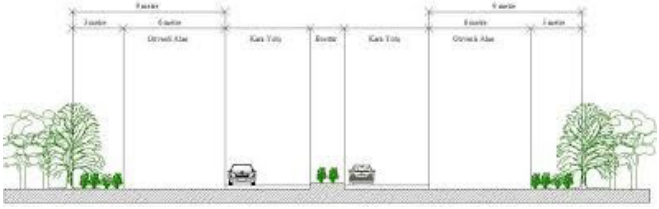


Figure 2. Design suggestion for road accident prevention [22].

Creating wind and snow cover:

- Plant covers to be formed to protect highways from wind should be positioned perpendicular to the wind direction and parallel to the leveling curves. Plant covers for snow should be installed 20-25m away from the road. Care should be taken that these plant covers do not block the landscape beauties that fall within the field of view from the road.
- Considering the problems of salt application in areas where there is a danger of snow and icing, salt-resistant plant species should be brought in these sections (Figure 3) [23].
-

Figure 3. Wind cut plant design [21]

Signalization (optical routing)

In undulating terrain with limited visibility, the driver will be warned by the appropriate afforestation method, for example the road route after a ramp (Figure 4).

Figure 4. Divided refuge strip on a highway in Huizhou [24].

III CONCLUSION AND SUGGESTIONS

In our country, where a large part of passenger and freight transportation is carried out by road, large shares are allocated from the budget for the construction and maintenance of roads. For this reason, in order not to disturb the ecology, the road constructions should be carried out where the road and the environment harmony are provided in the best way. With this study, it is aimed that highways are accepted as an important element of the landscape and in this context, the visual values of a highway environment are planned in a way that allows road users to make a safe and enjoyable journey. In addition, highway landscape planning should be handled from the routing stage, not after the road is opened to traffic.

From the stage of route determination and project planning, highways need the coordinated work of road engineers, city and regional planners, traffic engineers and landscape architects. In this process, the goal should be to provide the driver and passengers with a comfortable and safe journey with as little damage to nature as possible. In order to achieve this, plant material, which is separated from other materials with its effective functions in terms of engineering, traffic safety, ecological, visual and aesthetic aspects, should be used frequently on our highways.

REFERENCES

[1] Çalış, EA.2016. Monoray Ulaşım Sisteminin Özellikleri Ve Diğer Kentiçi Ulaşım Araçları İle Karşılaştırılması. İstanbul Teknik Üniversitesi, Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi.

[2] Altera, AZA, Bayraktar OY, Cetin M. 2019. Advanced Road Materials in Highway Infrastructure and Features. Kastamonu University, Journal of Engineering and Sciences. KUJES 5(1):36-42, 2019.

[3] Çetin İ.Z., Cesur A., Keskin R., & Akarsu H. (2018) Bazı Peyzaj Bitkilerinde Klorofil Miktarının Değişimi: Samsun Örneği, Kastamonu University Journal of Engineering and Sciences 4(1):1-10, 2018, https://dergipark.org.tr/download/article-file/609695

[4] Shirley, C., & Winston, C. (2004). Firm inventory behavior and the returns from highway infrastructure investments. Journal of Urban Economics, 55(2), 398-415

[5] Korkut, A. Kiper, T., Üstün Topal, T. 2017. Kentsel Peyzaj Tasarımda Ekolojik Yaklaşımlar. Artium, Vol. 5, No.1, 14-26, 2017.

[6] Selimoğlu, B. (1994). Ülkemiz Otoyollarında Çevre, Düzenleme İlkelerinin Belirlenmesi Üzerine Bir Araştırma. Ankara Üniversitesi Fen Bilimleri Enstitüsü Peyzaj Mimarlığı Anabilim Dalı Doktora Tezi, Ankara.

[7] Ertekin, M. ve Çorbacı, Ö.L. (2010). Karayollarında Peyzaj Planlama ve Bitkilendirme Çalışmaları. Ecological Life Sciences, 5(2): 106-125

[8] Koç, N. (1979). Karayolları Ağaçlamasını İşlev ve Estetik Yararları, Teknik Yöntemleri. Peyzaj Mimarlığı Dergisi Karayolları Özel Sayısı, 1: 13-18.

[9] Sosyaler, H. (1973). Yol Boylarında Erozyon Kontrolü Emniyet ve Estetik. Karayolları Genel Müdürlüğü Bakım Dairesi Başkanlığı Bakım Fen Heyeti Müdürlüğü, Ankara.

[10] Çorbacı, Ö.L. ve Var, M. (2011). Bartın-Amasra Karayolunun Peyzaj Özelliklerinin Peyzaj Planlama Açısından İrdelenmesi vee Sorunların Giderilmesine Çeşitli Öneriler. Bartın Orman Fakültesi Dergisi, 13(20): 23-37.

[11] Saatçioğlu, F. (1960). En Önemli Silvikültürel ve Estetik Özellikleriyle Şehir Ve Yol Ağaçları. İstanbul Üniversitesi Orman Fakültesi Dergisi, B(10,2):1-24

[12] Akdoğan, G. (1967). Ankara-İstanbul Karayolu Güzergâhının Peyzaj Özeliklerinin Etüdü ile Peyzaj Planlaması Yönünden Ele Alınması Gereken Problemler. Karayolları Genel Müdürlüğü Yayınları No: 158, Ankara.

[13] Tanrıverdi, F. (1975). Karayolları Ağaçlandırma Rehberi. Karayolları Genel Müdürlüğü Matbaası, Ankara.

[14] Sezen, I.2018. Karayolu Peyzajı ve Manzara Yolları. Süleyman Demirel Üniversitesi Mimarlık Bilimleri ve Uygulamaları Dergisi Araştırma makalesi MBUD 2018, 3(1):54-65. e-ISSN: 2548-0170

[15] Umar, F. ve Yayla, N. 1993. Yol İnşaatı, İstanbul Teknik Üniversitesi Rektörlüğü Yayınları, No:1541, İstanbul.

[16] Itır, O. 2013. Kayseri - Melikgazi - Sarımsaklı Köyü Arazi Toplulaştırma Projesi Ve Toplulaştırma Sonrasında Yol – Parsel Durumlarının Analizi. Ankara Üniversitesi Fen Bilimleri Enstitüsü Yüksek Lisans Tezi. Ankara

[17] Koç, N., Şahin, Ş., (1999). Kırsal peyzaj planlaması. Ankara üniversitesi Ziraat Fakültesi Yayın No: 1509(463), Ankara.

[18] Dağıstanlıoğlu, C. (2007). Isparta-Eğirdir Karayolunun Peyzaj Planlama İlkeleri Açısından İncelenmesi. T.C. Selçuk Üniversitesi Fen Bilimleri Enstitüsü Peyzaj Mimarlığı Ana Bilim Dalı. Konya.

[19] Ay, EA.2012. Havaalanı Kent Bağlantılarında Karayolu Peyzaj Düzenlemesi. İstanbul Teknik Üniversitesi, Fen Bilimleri Enstitüsü, Yüksek Lisans Tezi. İstanbul

[20] Yurt TD.2009. Ankara-İstanbul Otoyolu Bitkilendirme Çalışmalarının İrdelenmesi, Akıncı Gişeleri Çeltikçi Kavşağı Örneği. Ankara üniversitesi fen bilimleri enstitüsü. Yüksek lisans tezi. Ankara

[21] Anonymous, 2020. https://peyzax.com/ruzgar-perdesi-agaclari(Accessed date :21.09.2020)

[22] Yaşar MO. 2018. Kastamonu – Araç Kara Yolunun Peyzaj Planlama İlkeleri Açısından Değerlendirilmesi. T.C. Kastamonu Üniversitesi Fen Bilimleri Enstitüsü. Yüksek Lisans Tezi. Kastamonu

[23] Altınçekiç, SÇ, Altınçekiç, H.1999. Karayolları Peyzaj Düzenleme Çalışmalarında Bitkilendirme Esasları. İstanbul Üniv. Orman Fak. Dergisi, Seri B, Cilt 49. Sayı1-2-3-4, 1999.

[24] Anonymous, 2020. https://commons.wikimedia.org/wiki/File:Divided_median_strip_on_a_highway_in_Huizhou.jpg (Accessed date. 01.10.202)

Biomedical Engineering

Computer Engineering

Electrical and Electronics Engineering

Mechanical Engineering

Mechatronic Engineering

Metallurgical and Materials Engineering

Civil Engineering

ICENTE'20

INTERNATIONAL CONFERENCE ON ENGINEERING TECHNOLOGIES

November 19-21, 2020

Konya/TURKEY

icente.selcuk.edu.tr