INNOVATIONS AND TECHNOLOGIES IN ENGINEERING

EDITORS: PROF. DR. SAKIR TASDEMIR,
ASSOC. PROF. DR. ILKER ALI OZKAN

INNOVATIONS AND TECHNOLOGIES IN ENGINEERING

EDITORS
PROF. DR. SAKIR TASDEMIR
ASSOC. PROF. DR. ILKER ALI OZKAN

EĞİTİM
yayınevi

TABLE OF CONTENTS

A REVIEW OF CNN APPLICATIONS IN MEDICAL IMAGE ANALYSIS

Kübra UYAR[1]

INTRODUCTION

Artificial intelligence (AI) is a field of study that deals with the use of computers to mimic human cognitive functions. AI is accomplished when machines "intelligently" perform tasks based on algorithms. Machine learning (ML) is a branch of AI that focuses on the use of data and algorithms to mimic the way humans learn, gradually increasing its accuracy. Deep Learning which takes its name from deep neural networks does not require any human intervention; it is an ML technique that uses algorithms and large datasets to find patterns and generate outputs and responses. DL uses a "programmable neural network" that allows machines to make the right decisions without the help of humans.

DL algorithms used in medical image analysis have become a reliable and robust tool. A dramatically huge increase in the amount of medical data directed the researchers to DL approaches. Thanks to mathematical models and hardware and software devices, DL has taken a wide place in this field. Deep neural network models are used for the learning process to extract features from various data such as images, sound, and text. DL has attracted the interest of researchers and has shown satisfactory results in different applications. DL methods are Convolutional Neural Network (CNN), Fully Convolutional Network (FCN), Recurrent Neural Network (RNN), Auto Encoders, and Restricted Boltzmann Machines (RBM).

Medical imaging techniques correspond to looking inside the body and alongside laboratory tests are one of the most common types of the medical treatment process. Scanning of different organs and anatomical structures is provided by several medical modalities. X-rays, computed tomography, magnetic resonance imaging (MRI), mammography, positron emission tomography (PET), retina photography, and ultrasonography are among medical imaging techniques. Some modalities are used to capture specific organs such as retina and dermoscopic photography while others can analyze multiple organs at once such as CT and MRI.

The primary aim of medical image analysis is to categorize images into multiple classes or discover and detect important information inside of the image. Since it supports medical experts in accurately diagnosing a disease, medical image analysis is crucial. DL methods provide earlier diagnosis of diseases, diminish the workload

1   Selcuk University, Faculty of Technology, Computer Engineering, Konya, Turkey.

of the medical experts, and prevent the conflict of opinions among doctors.

This paper reviews the CNN model and its applications in medical imaging such as image classification, localization, detection, and segmentation.

The rest of this paper is organized as follows: Section 2 explains the basic structure of CNN model with layers. Medical image applications are detailed in Section 3. Finally, Section 4 summarizes the conclusions of the study.

CNN ARCHITECTURE

CNNs are a class of artifical neural network commonly used in computer vision applications. CNNs contain deep layers that transform the inputs into small fields with convolution filters. CNN consists of the input layer, convolution layer, activation layer, pooling layer, fully-connected layer, and classification layer as shown in Figure 1.
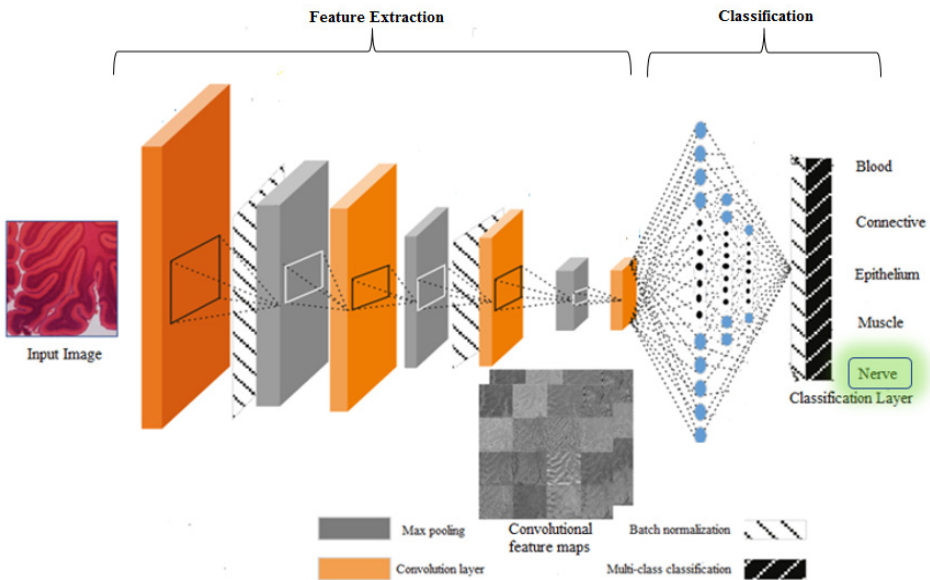
Figure 1. CNN architecture.

The details of the CNN layer are explained as follows:

The convolution layer enables the extraction of the features of the input image. Feature maps are created by circulating convolution filters that extract both low and high-level features on the input image. Feature maps are maps in which different features of the input data are extracted and the complexity of the model. Figure 2 shows an example of a convolution operation applied to the image with a 3x3 convolution filter.

Image 3x3 convolution filter Feature map (convolved feature)
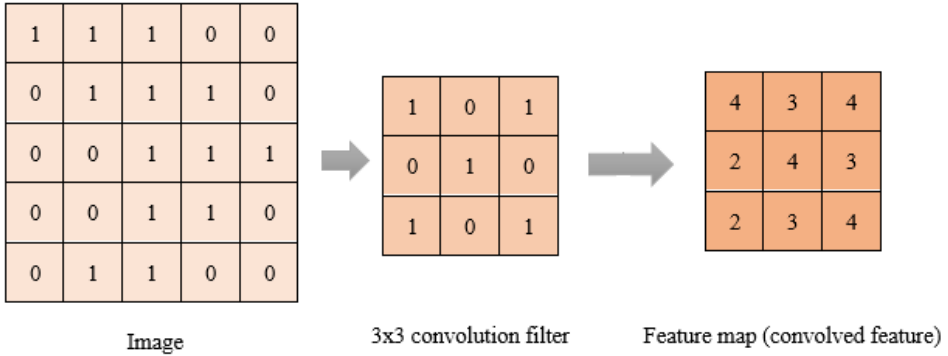
Figure 2. Convolution operation a 5x5 image with a 3x3 convolution filter.

*The pooling laye*r reduces the spatial size of convolved features. Dimensional reduction in convolved features provides a decrease in the number of parameters and dimensions of the network. Although the reduction in size causes information loss, it is important that the computational load is reduced and the memorization in the network is minimized. Furthermore, this layer allows extracting dominant features Maximum and average pooling are two types of pooling operations. Figure 3 demonstrates the example of a 3x3 maximum pooling operation over 5x5 convolved features.

Convolved feature 3x3 maximum pooling
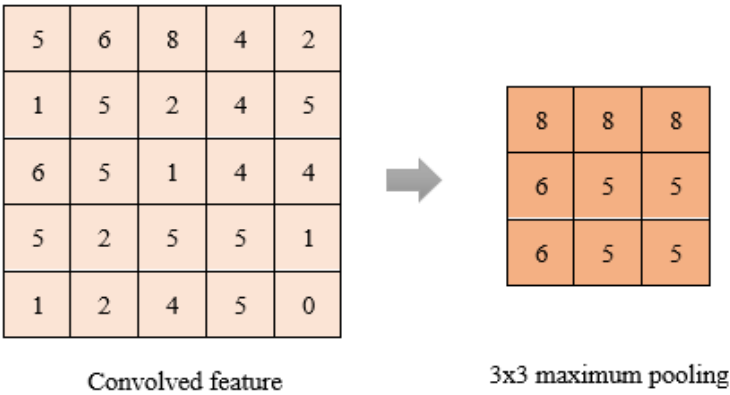
Figure 3. 3x3 pooling over 5x5 convolved feature.

The activation layer determines how the outputs produced after the convolution process will change. Since nonlinear activation functions increase the complexity of the model, activation functions such as sigmoid, tangent hyperbolic, and ReLU are used. Activation functions also ensure that the outputs stay within a certain value range. ReLU activation function is generally preferred in CNN. This function outputs 0 for negative values, while it produces the same output for positive values. Equation 1 expresses the mathematical representation of the ReLU activation function.

$$f(x) = max(0,x) \tag{1}$$

The fully-connected layer is the last layer of the feature extraction part of CNN. This layer consists of a series of fully-connected layers. In this layer, all the neurons of the previous layer are connected to every neuron of the current layer. To prevent overfitting during training, dropout can be applied as shown in Figure 4.

(a) Fully-connected layers   (b) After applying dropout
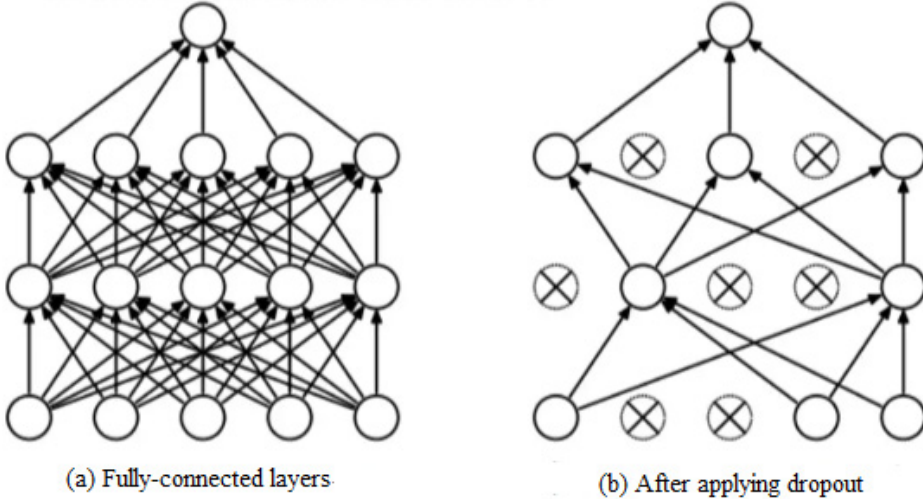
Figure 4. Applying dropout in fully-connected layers [1].

The classification layer computes the cross-entropy loss for classification task.

Data augmentation methods [2-4] have a positive effect on the performance of DL netwoks. In case of insufficient data, data augmentation methods such as reflection, rotation, translation, cropping, and flipping can be used. It improves the performance of deep models and balances the datasets.

MEDICAL IMAGE ANALYSIS TASKS

Wide variations in pathology and the potential tiredness of medical experts result in situations that are prone to human error. However, computer-aided decision support systems developed using DL methods provide a significant contribution to medical professionals. Medical data obtained with different imaging techniques can be analyzed quickly with DL methods. Figure 5 shows some medical imaging applications using various medical imaging techniques.

Some medical image applications can be listed as follows: brain lesions segmentation, prostate segmentation, mammographic mass classification, diabetic retinopathy classification, nodule classification, skin lesion classification, breast cancer metastases detection, and bone suppression.
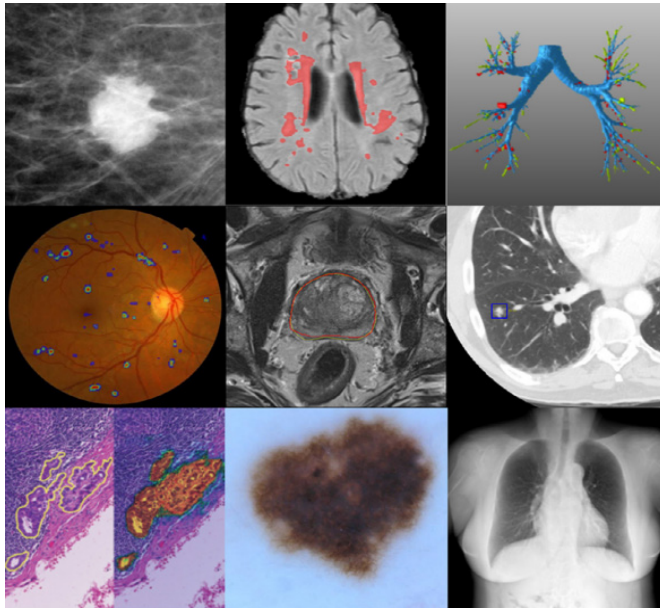
Figure 5. Some medical imaging applications [5].

Brain, eye, chest, breast, cardiac, abdomen, musculoskeletal, heart, kidney, and liver are among mostly analyzed images using DL methods. The use of medical imaging in DL can be divided into for main tasks: classification, localization, detection, and segmentation as illustrated in Figure 6.

Classification

Classification of medical images is one of the most important problems in image recognition and it aims to classify medical images into various categories to help medical experts. Extracting effective features from the image and building models that classify the image dataset using features are the two main steps for medical image classification. The classification model extracts the features of the input image and assigns labels using those features.

There are both traditional and dep model methods to solve image classification problems. Support vector machines [6, 7], random forest [8] and color and texture [9-12] are among traditional methods. Deep models to classify medical images contain [13-17]. In traditional methods color, texture, and shape features that are gained through experience are needed to classify images. For deep models, they extract the features of the input image during network training. In the experimental studies, pre-trained CNN models can be used for the classification task.
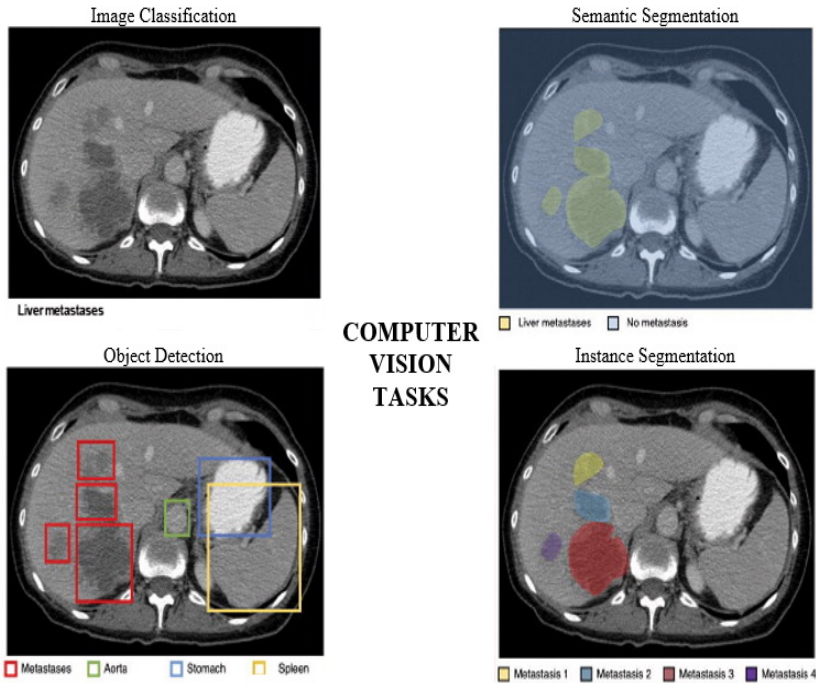
Figure 6. Computer vision tasks[18].

Analyzing lung nodules in breast X-ray images, classifying histological and histopathological images, and detecting Alzheimer's disease or diabetic retinopathy are some medical image classification problems.

Localization

Localization is carried out by classifying pixels depending on image information in their neighborhood. Prediction of the object in the image, bounding box drawing, and labeling are the main steps of localization. There are some deep method solutions for medical image localization and orientation detection [19-22]. Localization of the liver, heart, kidney, and spleen are among localization applications.

Detection

Detection is the detection of instances of semantic objects of a particular class found in images. Object detection provides the coordinates of the detected object on the image. With the founded coordinates, the area where the object will be enclosed with a frame is also determined. Detection of lesions in a scan (e.g., MRI, CT) is important both for the patient and the medical expert. Region-based CNN models such as R-CNN, Fast R-CNN, and Faster R-CNN are widely used for the detection of abnormal parts or specified organs in medical images.

There are some deep method solutions for medical image detection [23-26]. Detection of abnormal lymph nodes, detection of cancerous lung tissues, and detection of lesions are among the detection applications.

Segmentation

Segmentation is used to specify the classes of objects in the image as well as their boundaries with masks rather than bounding boxes. There are two different segmentation techniques: semantic segmentation and instance segmentation. Semantic segmentation detects all objects in an image at the pixel level and extracts regions with objects belonging to different classes. Thus, all objects in the image are grouped separately. In instance segmentation, each pixel is assigned a certain class, but two different objects in the same class have different colors from each other. The segmentation model divides the image into many regions. U-Net model and its variations are mostly used as deep models for segmentation problems [27].

Medical image segmentation difficult task because of various challenges such as type of pathology and biological variations. There are some deep-method solutions for medical image segmentation [28-33].

Automatic segmentation of the organs such as the liver, prostate, or brain is significantly important for the planning of the surgery and preventing surgical errors.

CONCLUSIONS

DL is one of the most popular methods used in medical imaging analysis. DL, which has algorithms with different network architectures, is widely used in the field of health, especially in medical images. These methods are frequently used in areas such as early diagnosis of diseases, early treatment, reducing the workload of medical experts, and different expert opinions. Classification, localization, detection, and segmentation are well-known tasks in medical image processing.

Images obtained with different medical imaging techniques lead to the formation of large data sets. DL methods are preferred instead of traditional methods in the analysis of large-scale medical data. It can be concluded that there is no specific algorithm or methodology that works well for all medical data. The characteristics of the specific data should be evaluated and then should be experimented with related DL methods.

REFERENCES

[1] Srivastava, N., et al., *Dropout: a simple way to prevent neural networks from overfitting.* The journal of machine learning research, 2014. 15(1): p. 1929-1958.

[2] Perez, L. and J. Wang, *The effectiveness of data augmentation in image classification using deep learning.* arXiv preprint arXiv:1712.04621, 2017.

[3] Xu, Y., et al., *Improved relation classification by deep recurrent neural networks with data augmentation.* arXiv preprint arXiv:1601.03651, 2016.

[4] Wong, S.C., et al. *Understanding data augmentation for classification: when to warp?* in *2016 international conference on digital image computing: techniques and applications (DICTA)*. 2016. IEEE.

[5] Litjens, G., et al., *A survey on deep learning in medical image analysis.* Medical Image Analysis, 2017. 42: p. 60-88.

[6] Zhang, Y.D., et al., *Magnetic resonance brain image classification based on weighted-type fractional Fourier transform and nonparallel support vector machine.* International Journal of Imaging Systems and Technology, 2015. 25(4): p. 317-327.

[7] Zhang, Y., et al., *Magnetic resonance brain image classification via stationary wavelet transform and generalized eigenvalue proximal support vector machine.* Journal of Medical Imaging and Health Informatics, 2015. **5**(7): p. 1395-1403.

[8] Ramírez, J., et al., *Computer aided diagnosis system for the Alzheimer's disease based on partial least squares and random forest SPECT image classification.* Neuroscience letters, 2010. 472(2): p. 99-103.

[9] Barata, C., et al., *Two systems for the detection of melanomas in dermoscopy images using texture and color features.* IEEE systems Journal, 2013. 8(3): p. 965-979.

[10] Riaz, F., et al., *Content-adaptive region-based color texture descriptors for medical images.* IEEE journal of biomedical and health informatics, 2015. 21(1): p. 162-171.

[11] Iyatomi, H., et al., *An improved internet-based melanoma screening system with dermatologist-like tumor area extraction algorithm.* Computerized Medical Imaging and Graphics, 2008. 32(7): p. 566-579.

[12] Stoecker, W.V., et al., *Detection of granularity in dermoscopy images of malignant melanoma using color and texture features.* Computerized Medical Imaging and Graphics, 2011. 35(2): p. 144-147.

[13] Li, Q., et al. *Medical image classification with convolutional neural network.* in *2014 13th international conference on control automation robotics & vision (ICARCV)*. 2014. IEEE.

[14] Bar, Y., et al. *Deep learning with non-medical training used for chest pathology identification.* in *Medical Imaging 2015: Computer-Aided Diagnosis*. 2015. SPIE.

[15] Shin, H.-C., et al., *Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning.* IEEE transactions on medical imaging, 2016. 35(5): p. 1285-1298.

[16] Chan, T.-H., et al., *PCANet: A simple deep learning baseline for image classification?* IEEE transactions on image processing, 2015. 24(12): p. 5017-5032.

[17] Ciregan, D., U. Meier, and J. Schmidhuber. *Multi-column deep neural networks for image classification.* in *2012 IEEE Conference on Computer Vision and Pattern Recognition*. 2012.

[18] Cheng, P.M., et al., *Deep learning: an update for radiologists.* Radiographics, 2021. 41(5): p. 1427-1445.

[19] Zhao, Y., et al., *Deep learning solution for medical image localization and orientation detection.* Medical Image Analysis, 2022. 81: p. 102529.

[20] Roth, H.R., et al. *Anatomy-specific classification of medical images using deep convolutional nets.* in *2015 IEEE 12th international symposium on biomedical imaging (ISBI)*. 2015. IEEE.

[21] Shin, H.-C., et al., *Stacked autoencoders for unsupervised feature learning and multiple organ detection in a pilot study using 4D patient data.* IEEE transactions on pattern analysis and machine intelligence, 2012. 35(8): p. 1930-1943.

[22] Liao, H., A. Mesfin, and J. Luo, *Joint Vertebrae Identification and Localization in Spinal CT Images by Combining Short- and Long-Range Contextual Information.* IEEE Transactions on Medical Imaging, 2018. 37(5): p. 1266-1275.

[23] Yildirim, O., et al., *Automated detection of diabetic subject using pre-trained 2D-CNN models with frequency spectrum images extracted from heart rate signals.* Computers in Biology and Medicine, 2019. 113: p. 103387.

[24] Srivastava, S., et al., *Comparative analysis of deep learning image detection algorithms.* Journal of Big Data, 2021. 8(1): p. 66.

[25] Ganatra, N. *A Comprehensive Study of Applying Object Detection Methods for Medical Image Analysis*. in *2021 8th International Conference on Computing for Sustainable Global Development (INDIACom)*. 2021.

[26] Valverde, C., et al., *Automated detection of diabetic retinopathy in retinal images.* Indian journal of ophthalmology, 2016. 64(1): p. 26.

[27] Ronneberger, O., P. Fischer, and T. Brox. *U-net: Convolutional networks for biomedical image segmentation*. in *International Conference on Medical image computing and computer-assisted intervention*. 2015. Springer.

[28] Akkus, Z., et al., *Deep Learning for Brain MRI Segmentation: State of the Art and Future Directions.* Journal of Digital Imaging, 2017. 30(4): p. 449-459.

[29] Wang, R., et al., *Medical image segmentation using deep learning: A survey.* IET Image Processing, 2022. 16(5): p. 1243-1267.

[30] Brosch, T., et al., *Deep 3D convolutional encoder networks with shortcuts for multiscale feature integration applied to multiple sclerosis lesion segmentation.* IEEE transactions on medical imaging, 2016. 35(5): p. 1229-1239.

[31] Li, X., et al., *3D multi-scale FCN with random modality voxel dropout learning for Intervertebral Disc Localization and Segmentation from Multi-modality MR Images.* Medical Image Analysis, 2018. 45: p. 41-54.

[32] Ronneberger, O., P. Fischer, and T. Brox. *U-Net: Convolutional Networks for Biomedical Image Segmentation*. in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. 2015. Cham: Springer International Publishing.

[33] Thomas, E., et al., *Multi-Res-Attention UNet: A CNN Model for the Segmentation of Focal Cortical Dysplasia Lesions from Magnetic Resonance Images.* IEEE Journal of Biomedical and Health Informatics, 2021. 25(5): p. 1724-1734.

ARTIFICIAL INTELLIGENCE APPLICATIONS IN FOOTBALL: A REVIEW

Mustafa AL-ASADI[1], Kübra UYAR[2]

INTRODUCTION

Football is the world's most popular sport, both in terms of players and spectators. Football's popularity has grown recently, making it an important part of the global economy. Machine learning (ML) is a branch of artificial intelligence (AI) that can be used to transform football statistical data into valuable information for coaches to use in analyzing opponents and making real-time decisions.

In this chapter, we will review three models of football clubs that rely on machine learning in their training programs, which are Benfica F.C. of the Portuguese Football League and the Seattle Reign football club of the U.S. women's football league, and theoretical studies related to this area will be reviewed in the literature review. Benfica F.C., In the last years, started using machine learning and technology to become one of the best clubs in the world by using a high-tech training simulator system called 360S, this system gives users access to training facilities. The system has helped Benfica win four domestic leagues over the past four years, and this system has also helped Benfica generate more than $350 million through the sale and transfer fees of players. The 360S lab is outfitted with canons that fire footballs at players and light-emitting diodes that provide ever-changing targets, along with a variety of sensors and computers as illustrated in Figure 1. By using the technology of tracking GPS, experts can supervise players' work rates in the training stadium. The trackers can keep track of a player's speed, distance covered during play, and pulse rate per game. All of this data can be used to forecast player injuries. Once the members of the team enter the campus, all their activity is recorded and stored in the "data repository" [1].

1   KTO Karatay University, Computer Engineering, Konya, Turkey, mustafa.alasadi@karatay.edu.tr
2   Selcuk University, Computer Engineering, Konya, Turkey, kubrauyar@selcuk.edu.tr

Figure 1. Target Wall L.E.D. of 360s System

The 360S relies on an LED wall, an automated "ball boy," and GPS tracking and uses fully customizable computer software to analyze performance and allow players to work on their skills [2].

Figure 2 is an example of working 360s in the analysis of quality sleep for the player based on his self-reported data and his sleep tracker. The figure show that the player is suffering from an insufficiency of sleep—both quantity and quality. As a result, the coaching staff might decide to switch player halfway through the game for a better-rested player [3].
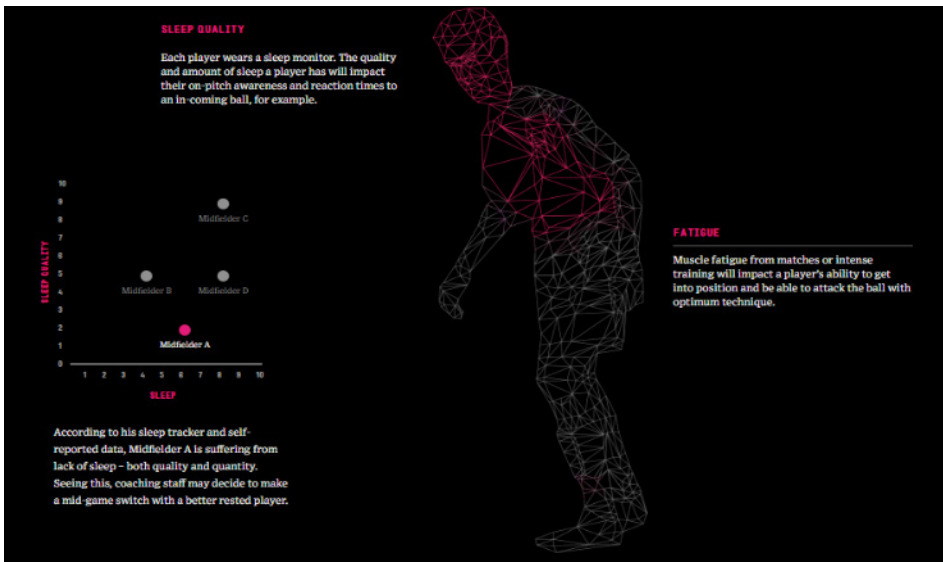
Figure 2. Analyze a Quality Sleep for Players in 360s System

In the middle of 2017, the Seattle Reign football club started using a new platform designed by POP digital marketing agency and Microsoft Corp., as shown in Figure 3. With the help of the Sports Performance Platform, sports teams can learn about how well their team and players are doing in order to come up with

useful insights that can affect how well they perform on the field. Using machine learning and predictive modeling, sports scientists can learn about athlete readiness, injury prevention, practice design and management, longitudinal player reports, and talent identification thanks to the platform. Additionally, teams can use the Sports Performance Platform to communicate important insights to team personnel like coaches, players, and management in order to guide decision-making that improves performance.

The platform uses sensors and surveys that athletes complete prior to, during, and after matches and training as well as wearable devices on the athletes to calculate their performance. The coaches can use this system to create reports that compare data over a wide range of timelines, such as comparing the first and second halves of a game. The platform would keep track of trends over time, such as how a player's performance was affected by how many hours of sleep they got before a match. This would help coaches better understand athletes and make adjustments to their training to improve performance. Data visualizations of the performance platform depend on the Microsoft Power BI reports. Further, the data housed by the system lives on the Microsoft Azure cloud computing platform [4].

Figure 3. Microsoft's Performance Platform for Seattle Reign F.C.

LITERATURE REVIEW

As a result of a literature review, ML has been used in five topics in football which are:

1- Predict football match result

2- Predict football players' injuries

3- Evaluate players and select the best players for the formation

4- Predict player skills, wages, and value

5- Football analytics

We explained the most important algorithms used in research and their efficiency and sources of data collection and the most important attributes affecting research.

Predict Football Match Result

A lot of efforts have been made to predict the results of the matches and to select the important variables in the prediction process. Prediction is critical in football to help club managers and coaches decide which matches and tournaments to enter. Businesses and gamblers have been trying to predict football game results for both tournaments and single matches. On the other hand, organized football gambling has developed into a growing industry and is now worth billions. As a result, there is much literature on match prediction models. In this section, we will present an overview of the most relevant research on match result prediction.

Hijmans et al. [5] proposed a learning algorithm through multiple data mining, which was analyzed and compared to determine the appropriate model for predicting the Dutch football team's matches. One single model is chosen based on the prediction results of the Nave Bayes model, a random tree model, and a k-nearest neighbor model. The random tree model, which had the most predictive power, is used to examine the results in greater depth.

Razali et al.[6] proposed a Bayesian network-based learning algorithm for predicting matches' outcomes in terms of win at home (H) or win away (A) or draw (D). For the three seasons 2011, 2012, and 2013, the English League is chosen. Prediction accuracy has been evaluated through K-fold cross-validation. The predicted accuracy of Bayesian networks is 75.09 percent.

Kınalıoğlu et al. [7] using ANN, SVM, and KNN algorithm techniques, we predicted the outcomes of 15 elimination rounds in the 2017 UEFA Champions League, including eight second rounds, four quarterfinals, a semi-final, and the final. "whoscored.com" provided the statistical data for seven seasons played between 2010 and 2016. Soccer statistics are compiled and used as training data by this host on a regular basis. The study concluded with a comparison of the various prediction methods' success rates.

Velcich [8] attempted to predict the outcomes of European league games using machine learning methods. The researcher calculated parameters for several machine learning algorithms using match statistics from Football-Data.co.uk: support vector machines, a random forest classifier, quadratic discriminant analysis, and polynomial regression

A logistic regression model designed to predict the English League's home or away wins for the 2015–2016 season was proposed by Prasetio [9]. Additionally, they utilized FIFA game data. The constructed model had a prediction accuracy of 69.5%.

Bo Shen [10] proposed a neural network-based learning algorithm for predicting soccer games' outcomes based on a variety of factors, including the skills of players and coaches, the effect of home ground on away ground, team tactics, and so on. Football manager (FM), a computer game, served as their data source. They were able to predict soccer football results with a test error of about 25%, proving that the neural network is the best model for the problem.

Wang et al. [11] proposed a learning algorithm that uses artificial neural networks to predict soccer matches' outcomes based on a variety of factors, such as the abilities of coaches and players; the home field away from the ground effect; team strategies; and so on.

Tax et al. [12] proposed a public data-based system for the Dutch Eredivisie to predict matches. A self-created public dataset containing match data from thirteen seasons of the Dutch Eredivisie was used for model training. A few mixes of dimensionality decrease strategies and grouping calculations have been tried on the common information preparing set in an organized manner. A multilayer perceptron classifier or Naive Bayes was used in conjunction with principal component analysis (PCA) to achieve the highest prediction accuracy on the public data feature set (with 15% variance).

Igiri [13] wanted to find out how well the support vector machine could predict match outcomes. A Gaussian combination kernel type is used to generate 79 support vectors after 100,000 iterations. 16 example football match results (data sets) were trained to predict 15 matches. The results revealed a prediction accuracy of 53.3%, which is relatively low. A study-developed SVM-based system is insufficient for this application domain.

Gomes et al. [14] proposed a decision support system to help customers of bookmakers increase profits on football-related bets. The goal of the project was to help people who bet on football win (away, home, or draw) and make more money.

Shin et al. [15] suggested using virtual games like FIFA to predict the outcome of a match. The real-time prediction was combined with the features of several players using logistic regression and linear support vector machines, with an accuracy predictor of 75% and a virtual predictor of 80%, respectively.

Arabzad et al. [16] proposed machine learning algorithms for predicting the outcome of one week's play in the Iranian football league for the 2013-2014 season using games from the previous seven leagues; The outcomes have demonstrated that neural networks can predict match outcomes.

By analyzing football games from various leagues, Yezus [17] sought to determine whether it is possible to accurately predict sports games' outcomes. There are two factors at play. To begin, carefully select the features that appear to be significant and evaluate their impact on the match's outcome. Second, by employing

techniques from machine learning like KNN, Random Forest, logistic regression, SVM, etc., He tried several times to build a model that could accurately predict the outcome of the match. However, it appears to be extremely challenging to succeed in this field. The study's overall success can be attributed to its 70% accuracy in predicting the outcome.

Moroney [18] suggested looking at football scores from football-data.co.uk to see if match facts, like goals, fouls, shots on target, and other things, were present. can predict how a match will end. The goal is to improve the skills learned in the course in areas like databases, programming, statistics, business analysis, and data mining. The construction of a database, statistical analysis in R, and machine learning in WEKA were used to carry out the analysis. Fouls, shots on goal, and other factors all have correlations, as the study demonstrated. and that match facts can be used to predict the game's outcome.

Igiri et al. [19] offered suggestions for data analyses; They use a tool called rapid miner and a process called Knowledge Discovery in Database to predict who will win the match. Both artificial neural networks and logistic regression are utilized as classifiers. The prediction of the match winner is made with an accuracy of 93%.

Ulmer et al. [20] proposed a hidden Markov model, Naive Bayes, Linear from stochastic gradient descent, Random Forest, and Support Vector Machine machine learning algorithm, for predicting English Premier League football match outcomes. To determine the most effective strategy, the accuracy of each model was calculated. After comparing all of the previous approaches, they discovered that support vector machines were the most effective, with a prediction accuracy of between 55 and 69 percent.

Owramipur et al. [21] suggested utilizing a Bayesian Network (BN) to forecast Barcelona Football Club matches. The Spanish football league's 2008-2009 season was the subject of the investigation. They discovered that the Bayesian Network (BN) can correctly predict the outcomes of upcoming football matches, with 92% accuracy.

Constantinou et al. [22] proposed utilizing a Bayesian organization model for the expectation Football result as per information and information, to anticipate English Chief Association (EPL) matches before they start, and exhibited benefit against all of the market chances, and contrasted and one more distributed football forecast models, pi-football it demonstrated outstandingly precise in expectation.

Hucaljuk et al. [23] are solutions to the problem of predicting football outcomes that are proposed by means of a machine learning model. Throughout the model's development, a number of tests were conducted to select the best attributes and classifications. This model's outcomes demonstrate a strong capacity for prediction.

Huang et al. [24] proposed a multilayer perceptron-based prediction model with a backpropagation learning rule. If the draw games are excluded, the prediction

accuracy based on the M.L.P. method can reach 76.9%. The multilayer perceptron (MLB) with backpropagation neural network learning is the foundation of the prediction system. The model's prediction accuracy was 76.9%.

A method that outperforms bookmakers' odds for predicting football matches has been proposed by Buursma [25]. The teams' previous performances serve as the foundation for the predictions for the games.

Van Gemert et al. [26] proposed a statistical model for Premier League football full-time scores. The statistical model takes into account the relationship between the home and away teams' respective goalscoring totals. The censored zero-inflated Poisson distribution and the censored Negative Binomial distribution were compared for the marginal distributions of the number of home and away goals. Additionally, the profitability of these models in comparison to bookmakers was examined.

Joseph et al. [27] suggested using a Bayesian network and machine learning algorithms to predict whether Tottenham Hotspur football club would win, lose, or draw a match. The Naive Bayesian Data-Driven Bayesian, MC4, K-nearest neighbor learner, and decision tree learner are all examples of machine learning techniques. The findings demonstrated that the predictive accuracy of the Bayesian network is superior to that of other methods.

Rotshtein et al. [28] suggested utilizing a fuzzy knowledge base based on the outcomes of previous matches. Based on previous results, they come to the conclusion that the match's development can be predicted.

The study of the literature in this area makes it abundantly clear that the majority of the machine learning algorithms used to predict matches' outcomes were limited to predicting wins, losses, and draws.

Predict Football Players' Injuries

In football, injuries are a big problem. It is thought to be the main reason football players don't go to games or practice, as well as the cost of their rehabilitation. Consequently, there is research on football players' injury prediction. The most significant studies on using machine learning to predict player injuries are listed below:

Based on machine learning and G.P.S. measurements, Rossi et al. [29] proposed a multidimensional approach to injury prediction in professional football. They use G.P.S. technology to collect information about how much work players at a professional football club put into training each season. By providing a collection of football practitioners' relevant case studies, they demonstrate the accuracy and interpretability of their injury predictors.

Carey et al. [30] examined Australian football players and proposed a learning algorithm for predicting athlete ratings of perceived exertion (RPE). Over the course of a full season, 45 players' heart rates and accelerometers were used to collect the

data from the global positioning system. The study has demonstrated, through the use of machine learning, that Australian football players' RPE can be predicted. Classification and linear models performed worse than regression modeling.

A learning algorithm was proposed by Kampakis [31] to investigate the predictability of football injuries. Together with Tottenham Hotspur F.C., this work was completed; The Gaussian process model was used to predict injuries, while negative binomial and ordinal regression as well as Poisson were used to predict the recovery time of injuries. Three studies were carried out: predicting an intrinsic injury, predicting the recovery time of injuries, and predicting player injuries. Finally, a different kind of algorithm—supervised P.C.A., naive Bayes, random forests, SVM, ANN, Ridge Logistic Regression (RLR), and KNN—was used to solve the third problem—predicting intrinsic damage.

19 Australian League football players were monitored for a full season using (G.P.S.) units in training and preseason games by Ehrmann et al. [32] to examine the relationship between variables measured by G.P.S. in gameplay and training. During the season, noncontact soft tissue injuries were documented, which indicated an increase in training and game intensity that led to injuries.

Kampakis [33] attempted to determine whether information at the time of injury could be used to predict the injured player's recovery time. Additionally, he employed three machine learning techniques—gothic processes, neural networks, and support vector machines. The data used in the tests came from Tottenham Hotspur F.C. The findings of the study show that this task can be done with some accuracy.

Venturelli et al. [34] employ a multivariate survival model (specifically, Cox regression) for youth players to investigate the factors that raise muscle pull risk. The study has demonstrated that the most significant factor is previous injuries. In addition, it demonstrated that having a larger stature increased the likelihood of muscle pull.

Brink et al. [35] aim to find out how measures to monitor stress and recovery, as well as their analysis, help to prevent injuries and illnesses in elite young football players. 53 elite football players between the ages of 15 and 18 participated in the study. To distinguish actual pressure, football players recorded preparing, span of the game, and assessment of the course of tension for two cutthroat periods through day to day preparing logs. Injury and illness data were gathered using FIFA's standard recording system, and M.R.A. was used to calculate OR and 95% C.I.s for injuries and illnesses, indicating that injuries are linked to physical stress.

According to the medical analysis of the players, machine learning algorithms were used to predict injuries in the players, particularly those involving the heart and muscles, as well as the durations of recovery from injuries.

Evaluate Players and Select Best Players for Formation

The final success of the selection of players and team formation is determined by how effectively the players are grouped together. To assist trainers in this field, highly structured models have been developed. The most significant studies on using machine learning to evaluate players and select the best players for formations are listed below:

Sathe et al. [36] proposed support vector machines, random forests, and naive Bayes as machine learning algorithms for feature selection in English premier league football.

Vroonen et al. [37] proposed a projection framework for Pertinent, roused by the CARMELO framework. APROPOS uses a historical dataset to search for player potential.

Soto-Valeroet et al. [38] suggested using principal component analysis (PCA) to describe football players' positions using a model-based Gaussian clustering method. 40 FIFA video features and 7705 players are used to test this model. There were three positions for the players to play. In addition, they discovered that the most distinct variable among various mixed player combinations was dribbling.

Asif et al. [39] was a one-of-a-kind instance in which a rating system for quantitatively assessing a player's performance was desired. This would ultimately empower forecasts deductions on different variables, like player execution or match results. Diverse sources were needed to gather player rating data; However, the methods used to collect such data are outlined in this Case Study.

The Player Performance Index (PPI), a statistic-based performance rating system for the Bundesliga, was proposed by Klaiber [40].

Abidin et al. [41] suggested a suitable research method that can be used in a Decision Support System (D.S.S.) study, particularly when the development of system artifacts is one of the research goals. The plan of the exploration strategy depended on the consummation of a football D.S.S. improvement that can assist with deciding the place of a player and the best group development to be utilized during a game. Researchers discovered, after reviewing the relevant literature for this study, that the conventional rainfall System Development Life Cycle (SDLC) approach and the Case Study approach must be combined to help structure the research task and phases and contribute to the achievement of the research goals.

Cotta et al. [42] suggested making use of FIFA game data as the dataset. They look at two recent topics that have been talked about a lot to justify their use and talk about possible applications.

Uzochukwu et al. [43] proposed a model that employs a neural network approach to identify these significant attributes for each player and divides them into four major categories for the purpose of player selection. These categories are the player's

technique, speed, physical status, and resistance. The outcome has demonstrated that a neural network can be used to select football team members.

Sarda et al. [44] suggested using a genetic algorithm to find the best solution to the team selection problem and team formation. They developed a model in this paper that combines a well-established quantitative approach with a few novel additions, such as characteristics pertaining to a player's individual and collective performance and the collaborative work of other players in the group.

Enefiok et al. [45] suggested a better system for managers to use in team selection that was made with fuzzy logic and artificial neural networks. The outcome demonstrates that the new decision support system has improved decision accuracy for player selection.

Tavana et al. [46] proposed a method for selecting the best football team formation in two steps: selecting the players first, then selecting the best formation. The players are evaluated using a hazy ranking in the first step, and the best performers are chosen for the team. A fuzzy inference system is used in the second phase to evaluate the various combinations of the selected players and determine the best team formations. This method improves the quality of the coaches' decisions and helps them solve decision-making issues. When evaluating players, coaches' judgments are essential; As a result, the coaches' cognitive abilities determine the model's effectiveness.

Kumar [47] tries to discover the hidden information that experts use to assign ratings to football players by attempting to classify them according to the performance of the essential player. The researcher used various Machine Learning algorithms and conducted three classification experiments. The best performance metric ratings prediction results had a mean absolute error of 0.17.

Bazmara et al. [48] proposed a KNN learning calculation to assess football gifts for legitimate positions thinking about player abilities. Real data is used to make the selection of players using the proposed method; Furthermore, the outcomes demonstrate this strategy's high efficiency.

The A.H.P. decision support system (D.S.S.) was proposed by Febianto [49] to support a player's ideal placement by selecting an appropriate player based on multiple criteria. D.S.S. would use A.H.P. as a model for various weighing in the selection process and assist the trainer in making the right choice. For related issues, literature, observation, and interviews are used as data collection methods. A data flow diagram (DFD) and an entity-relationship diagram (E.R.D.) are also used in an organized approach to data analysis models and procedures.

The literature review in this area yielded a few models for predicting the player's preferred team position, which were limited to attack, defense, and midfielder. In addition, a few algorithms are utilized for this purpose.

Predict Player Skills, Wages, and Value

Managers may be able to make better decisions about selling, buying, and renewing contracts by anticipating player skills, wages, and value. One of the most important ways to identify a player's talent is to also show their skill at passing, dribbling, and controlling the ball. Where these skills are the player's necessary technical skills, dribbling skill, in particular [50], is considered crucial to the match's outcome [51], and a previous study [38] found that dribbling was the most distinguishing variable among player skills. The most significant studies on using machine learning to predict player skills, wages, and value are listed below.

Using data on more than 15,000 players from the football simulation video game FIFA 2017, Dey [52] proposed a multilayer perceptron neural network to predict a soccer player's price. Experiments with various activation functions, layers and neurons, learning rate and decay, Nesterov momentum-based stochastic gradient descent, L2 regularization, and early stopping were used to improve the network. The trade-offs between the various aspects of neural network training are investigated simultaneously. The final model is accurate to within 6.32% of the actual price for any footballer, with a top-5 accuracy of 87.2% across 119 pricing categories.

Yaldo et al. [53] proposed a quantitative and objective method for evaluating the skills of football players to determine their wages. The experimental results show that the Pearson correlation between the players' actual and expected salaries is 0.77 (p 001) using data from 6082 players.

He et al. [54] demonstrated how extensive data sources and machine learning methods could be used to design players' market value and performance in La Liga.

from the study of this region's literature. We noticed that a number of models had been created to categorize players and national teams based on how well they performed. A model has been developed to estimate the player's market value based on his abilities and comprehension.

Football Analytics

To support training decisions, sports analysis makes use of quantitative data analysis of performance data. An analysis of performance data and research with a clear, practical purpose is sports analysis. Training programs, game development, player recruitment, and building strategies all benefit from sports analysis. Numerous commercial businesses have recently emerged to provide the sport's elite with data collection and analysis thanks to technological advancements [55]. The availability of match-by-match tracking data sparked a flurry of interest in football analytics. There is limited research in football analytics using Machine Learning methods that focuses on identifying formations and analyzing football game play [37]. There are many Games Examination Organizations giving tremendous volumes of information in measurable bundles and information perceptions. Prozone and Opta

Sports are the two businesses that have the greatest impact. The most significant studies on football analysis are listed below:

Wagenaar et al. [56] looked at how to use position data in football to use machine learning to predict opportunities for goals. For this issue, they suggest utilizing deep learning convolutional neural networks. The outcomes show that Google Net engineering is superior to any remaining strategies with an exactness of 67.1%.

Brooks et al. [57] suggested creating a novel player ranking system based on the value of passes completed. The relationship between pass locations and shot opportunities created during the possession is the foundation for this value. The model was built using data from the 2012-2013 La Liga season.

Sgro et al. [58] compare and contrast the technical performance profiles of the various teams during the 2016 European Football Championship. A k-implies bunch examination was for starters performed to recognize the nearby matches of that competition. After that, the Union of European Football Championship (UEFA) team-match statistics were gathered.

Horton et al. [59] suggested constructing a framework for rating each pass as "good," "OK," or "bad" based on the quality of the access and dividing passes made during a football game into these three categories. A list of tokens and player trajectories are required. The classification function was learned using supervised machine learning algorithms in the chosen method. Five classifiers were used in the experiments. They started by employing three distinct regularized cost functions in multinomial logistic regression. Second, the RUSBoost and Support Vector Machine algorithms were used as classifiers. On the pass labeling mission, they produced a classifier with 86 percent accuracy as a result.

Lasek et al. [60] gave an overview of how well the rating systems used by various football teams could predict outcomes. The FIFA ranking served as the primary standard. This system can perform better than the FIFA ranking, as demonstrated by their experiences.

Gedikli et al. [61] proposed the vision system ASPOGAMO, which can estimate the motion trajectories of football players captured on video. By employing model-based vision algorithms for player and camera estimation, the system achieves a high level of robustness.

Due to a lack of data, the literature review in this region revealed that few studies had been conducted to analyze sports, particularly football. Video games like FIFA Soccer, P.E.S., and Football Manager have been used to collect some sports analytics data.

CONCLUSIONS

Through their role in converting football statistics into useful information that assists teams, coaches, and athletes in analyzing opponents and making better decisions in real time, the study has demonstrated that machine learning plays a significant role in football. There have been a variety of machine learning methods utilized in this field.

This study's most significant findings are as follows:

1- Predicting football match outcomes has been the primary focus of previous research using machine learning techniques in football analytics. As a result, we recommend that developers investigate other topics like predicting injuries to football players or the best team formation.

2- Given that previous research has demonstrated that data collected from video games like FIFA, PES, and Football Manager can improve the quality of predictions, we recommend that game developers use these games as a data source.

3- Previous research has demonstrated that the Structured System Analysis and Design Methodology (SSDM) and the Object-Oriented Methodology (OOM) are the two primary approaches to the construction of prediction system models. For this reason, we recommend that web-based system developers employ SSDM.

REFERENCES

[1] PETTIT, H. How Benfica became one of the best football teams in the world by combining technology and data science. 2017 2/1/2021]; Available from https://www.dailymail.co.uk/sciencetech/article-4544900/How-world-s-best-football-clubs-use-data.html.

[2] Monaghan, M. High-tech training simulator set to be introduced to the U.A.E. 2017 09/24/2018]; Available from: http://sport360.com/article/other/31906/high-tech-training-simulator-set-be-introduced-uae/.

[3] wired. Infographic: How S.L. Benfica's football campus is powered by data. 2017 09/24/2018]; Available from: https://www.wired.co.uk/partnerships/infographic-sl-benfica.

[4] Day, M. Microsoft scrambles to become a player in the new game of sports analytics. 2017 09/24/2018]; Available from: https://www.seattletimes.com/business/microsoft/microsoft-scrambles-to-become-player-in-new-game-of-sports-analytics/.

[5] Hijmans, A. and S. Bhulai, Dutch football prediction using machine learning classifiers. 2017.

[6] Razali, N., et al. Using Bayesian networks to predict the outcomes of football matches in the English Premier League (EPL). In I.O.P. Conference Series: Materials Science and Engineering. 2017. I.O.P. Publishing.

[7] Kınalıoğlu, İ.H., C. Kuş, and İ. Kınacı. Data Mining and Artificial Neural Network Methods for Predicting the 2017 Uefa Champions League. in 3 rd INTERNATIONAL RESEARCHERS, STATISTICIANS AND YOUNG STATISTICIANS CONGRESS 24-26 MAY 2017 SELÇUK UNIVERSITY. 2017.

[8] Velcich, K., Predicting Soccer Match Results. International Journal of Smart Home, 2017. 9(7): p. 231-240.

[9] Prasetio, D. Using logistic regression to predict football match outcomes. In Advanced Informatics: Concepts, Theory And Application (ICAICTA), 2016 International Conference On. 2016. IEEE.

[10] Shen, B., Result Prediction for football Games, in CS221 Project Final Report. 2016: Stanford University

[11] Wang, S., et al., Short-term wind power forecasting using a neural network with back propagation. International Journal of Smart Home, 2015. 9(7): p. 231-240.

[12] Tax, N. and Y. Joustra, Predicting the Dutch football competition using public data: A machine learning approach. Transactions on Knowledge and Data Engineering, 2015. 10(10): p. 1-13.

[13] Igiri, C.P., Support Vector Machine-Based Football Match Result Prediction System. IOSR Journal of Computer Engineering (IOSR-JCE), 2015. 17(3): p. 21-26.

[14] Gomes, J., F. Portela, and M.F. Santos. System for making decisions and predicting the outcome of football games. The 19th International Conference on Circuits, Computers, Systems, Communications and Computers-Intelligent Systems and Applications Special Sessions. Series. 2015.

[15] Shin, J. and R. Gasparyan, A novel way to fotball Match Prediction. Stanford University: Department of Computer Science, 2014.

[16] Arabzad, S.M., et al., Prediction of football match outcomes using artificial neural networks; the case of Iran Pro League. Journal of Applied Research on Industrial Engineering, 2014. 1(3): p. 159-179.

[17] Yezus, A., Using machine learning to predict soccer matches' outcomes. Saint-Petersburg University, 2014.

[18] Moroney, K., Foreseeing match results through game occasions " 2014, national college of Ireland.

[19] Igiri, C.P. and E.O. Nwachukwu, An improved football match result prediction system. IOSR Journal of Engineering (IOSRJEN), 2014. 4: p. 12-20.

[20] Ulmer, B., M. Fernandez, and M. Peterson, Anticipating Soccer Match Results in the English Chief Association. 2013, Ph. D. dissertation.

[21] Owramipur, F., P. Eskandarian, and F.S. Mozneb, Prediction of the Barcelona-Barcelona match's outcome using a Bayesian network in the Spanish League. International Journal of Computer Theory and Engineering, 2013. 5(5): p. 812.

[22] Constantinou, A.C., N.E. Fenton, and M. Neil, pi-football: A Bayesian network model for predicting the outcomes of Association Football games. Knowledge-Based Systems, 2012. 36: p. 322-339.

[23] Hucaljuk, J. and A. Rakipović. Using techniques from machine learning to predict football scores. in MIPRO, 2011 Proceedings of the 34th International Convention. 2011. IEEE.

[24] Huang, K.-Y. and W.-L. Chang. Prediction of the football match at the 2006 World Cup using a neural network method. in Neural Networks (IJCNN), The 2010 International Joint Conference on. 2010. IEEE.

[25] Buursma, D. Predicting sports events from past results. in 14th Twente Student Conference on IT. 2010.

[26] Van Gemert, D. and J. van Ophem, Modelling the Scores of Premier League Football Matches. Economerics, 2010. 18: p. 67.

[27] Joseph, A., N.E. Fenton, and M. Neil, Predicting football results using Bayesian nets and other machine learning techniques. Knowledge-Based Systems, 2006. 19(7): p. 544-553.

[28] Rotshtein, A.P., M. Posner, and A. Rakityanskaya, Football predictions based on a fuzzy model with genetic and neural tuning. Cybernetics and Systems Analysis, 2005. 41(4): p. 619-630.

[29] Rossi, A., et al., Effective injury prediction in professional soccer with G.P.S. data and machine learning. arXiv preprint arXiv:1705.08079, 2017.

[30] Carey, D.L., et al., Predicting ratings of perceived exertion in Australian football players: methods for live estimation. International Journal of Computer Science in Sport, 2016. 15(2): p. 64-77.

[31] Kampakis, S., Predictive modelling of football injuries. arXiv preprint arXiv:1609.07480, 2016.

[32] Ehrmann, F.E., et al., G.P.S. and injury prevention in professional soccer. The Journal of Strength & Conditioning Research, 2016. 30(2): p. 360-367.

[33] Kampakis, S., Comparison of machine learning methods for predicting the recovery time of professional football players after an undiagnosed injury. 2011.

[34] Venturelli, M., et al., Injury risk factors in young soccer players detected by a multivariate survival model. Journal of science and medicine in sport, 2011. 14(4): p. 293-298.

[35] Brink, M.S., et al., Monitoring stress and recovery: new insights for the prevention of injuries and illnesses in elite youth soccer players. British journal of sports medicine, 2010. 44(11): p. 809-815.

[36] Sathe, S., et al., Predictive Analysis of Premier League Using Machine Learning. International Journal of Innovative Research in Computer and Communication Engineering, 2017. 5(3).

[37] Vroonen, R., et al. Predicting the potential of professional soccer players. in Machine Learning and Data Mining for Sports Analytics ECML/PKDD 2017 workshop. 2017.

[38] Soto-Valero, C., A Gaussian mixture clustering model for characterizing football players using the E.A. Sports' FIFA video game system. RICYDE. Revista Internacional de Ciencias del Deporte, 2017. 13(49).

[39] Asif, R., et al., Football (Soccer) Analytics: A Case Study on the Availability and Limitations of Data for Football Analytics Research. International Journal of Computer Science and Information Security, 2016. 14(11): p. 516.

[40] Klaiber, J.D., Soccer Player Performance Rating Systems for the German Bundesliga, in Department of Mathematical Statistics. 2016, Ghent University.

[41] Abidin, M.Z.Z., M.K.M. Nawawi, and M.M. Kasim. Research design of decision support system for team sport. in A.I.P. Conference Proceedings. 2016. A.I.P. Publishing.

[42] Cotta, L., et al. Using fifa soccer video game data for soccer analytics. in Workshop on Large Scale Sports Analytics. 2016.

[43] Uzochukwu, O. and P. Enyindah, A Machine Learning Application for Football Players' Selection. International Journal of Engineering Research & Technology (IJERT), 2015. 4(10).

[44] Sarda, V., P. Sakaria, and K. Deulkar, Football team selection using genetic algorithm. International Journal of Engineering and Technical Research, 2015. 3(2): p. 153-156.

[45] Enefiok, E., E. Nwachukwu, and E.E. Williams, An Improved Decision Support System for a Football Team Manager. International Journal of Engineering Research & Technology (IJERT), 2015. 4(5).

[46] Tavana, M., et al., A fuzzy inference system with application to player selection and team formation in multi-player sports. Sport Management Review, 2013. 16(1): p. 97-110.

[47] Kumar, G., Machine Learning for Soccer Analytics. 2013, Cambridge University Press, MSc thesis, K.U. Leuven.

[48] Bazmara, M. and S. Jafari, K Nearest Neighbor Algorithm for Finding Soccer Talent. Journal of Basic and Applied Scientific Research, 2013. 3(4): p. 981-986.

[49] Febianto, I., DECISION SUPPORT SYSTEM FOR IDEAL PLACEMENT OF PLAYERS POSITION STRATEGY IN FOOTBALL FORMATION. 2010, Universitas Komputer Indonesia.

[50] Reilly, T. and M. Holmes, A preliminary analysis of selected soccer skills. Physical Education Review, 1983. 6(1): p. 64-71.

[51] Huijgen, B.C., et al., Development of dribbling in talented youth soccer players aged 12–19 years: A longitudinal study. Journal of Sports Sciences, 2010. 28(7): p. 689-698.

[52] Dey, S., Pricing Football Players using Neural Networks. arXiv preprint arXiv:1711.05865, 2017.

[53] Yaldo, L. and L. Shamir, Computational Estimation of Football Player Wages. International Journal of Computer Science in Sport, 2017. 16(1): p. 18-38.

[54] He, M., R. Cachucho, and A. Knobbe. Football Player's Performance and Market Value. in Proceedings of the 2nd workshop of sports analytics, European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML PKDD), Porto, Portugal. 2015.

[55] Schulenkorf, N. and S. Frawley, Critical issues in global sport management. 2016: Taylor & Francis.

[56] Wagenaar, M., et al. Using Deep Convolutional Neural Networks to Predict Goal-scoring Opportunities in Soccer. in ICPRAM. 2017.

[57] Brooks, J., M. Kerr, and J. Guttag. Developing a data-driven player ranking in soccer using predictive model weights. in Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2016. A.C.M.

[58] Sgro, F. and M. Lipoma, Technical performance profiles in the European Football Championship 2016. Journal of Physical Education and Sport, 2016. 16(4): p. 1304.

[59] Horton, M., et al., Classification of passes in football matches using spatiotemporal data. arXiv preprint arXiv:1407.5093, 2014.

[60] Lasek, J., Z. Szlávik, and S. Bhulai, Ranking systems' ability to predict outcomes in association football. International Journal of Applied Pattern Recognition, 2013. 1(1): p. 27-46.

[61] Gedikli, S., et al. An adaptive vision system that uses different camera settings to follow soccer players. in Proceedings of the 5th International Conference on Computer Vision Systems (ICVS). 2007.

HYBRID GREY WOLF WITH HARRIS HAWKS OPTIMIZATION FOR BENCHMARK FUNCTIONS

Mustafa Serter UZER [1], Onur INAN [2]

INTRODUCTION

Solving some classes of problems using classic mathematical programming methodologies is rather challenging because many issues in machine learning and artificial intelligence frequently have a continuous, discrete, limited, or uncontrolled nature [1-3]. Because of their simplicity, ease of implementation, and problem-solving power, metaheuristic algorithms have been created and used for a range of problems as rival problem-solver. Furthermore, neither the mathematical characteristics nor the gradient information of the goal landscape is necessary for the basic operations of these systems. The main drawback of most metaheuristic algorithms, nevertheless, is that they frequently exhibit a delicate sensitivity to the adjustment of user-defined parameters. Another drawback issue is that the global optimum may not always be reached by the metaheuristic algorithms [1, 4]. Therefore, new metaheuristic optimization algorithms are constantly being developed to take advantage of their advantages and eliminate their disadvantages, or new hybrid algorithms are derived by combining these improved algorithms with other algorithms. Meta-heuristic techniques can be categorized into three groups: swarm-based, physics-based, and evolutionary-based [5]. Examples of swarm-based optimizations from the literature include: Butterfly Optimization Algorithm(BOA) [6], Harris Hawks Optimization (HHO) [1], Artificial Bee Colony (ABC) [7], Particle Swarm Optimization (PSO) [8], Whale Optimization Algorithm (WOA) [5], Optimal Foraging Algorithm (OFA) [9], Grey Wolf Optimization (GWO) [4].

The HHO algorithm is an optimization method that simulates the event of several hawks collaboratively attacking from different directions to confuse prey. According to the escape plans of the prey, hawks have developed various hunting strategies. As a result of the mathematical modeling of these various strategies, HHO was formed [1]. GWO is an optimization algorithm that models grey wolves' strategies to capture prey in the wild, and Mirjalili et al. developed it [4]. It is based on cooperative hunting techniques developed by grey wolves in accordance with their social hierarchies. Whale optimization algorithm (WOA), is put forth

1   Mustafa Serter Uzer, Dr, Department of Electronics and Automation, Ilgın Vocational School, Selcuk University, Turkey, ORCID No: https://orcid.org/0000-0002-8829-5987
2   Onur Inan, Dr, Department of Computer Engineering, Faculty of Technology, Selcuk University, Turkey, ORCID No: https://orcid.org/0000-0003-4573-7025

by Mirjalili and Lewis in 2016 [5]. The techniques utilized by humpback whales, which create bubbles and squeeze their meal into a bubble spiral, are served as the basis for WOA. Air bubbles help humpback whales gather their prey, which often comprises small fish communities then they rise to the top and move to focus on the target species in a smaller area. Encircling prey, launching a bubble-net attack, and looking for prey are the stages of WOA. A hybrid strategy known as HGWWO is presented in [10] and is based on the Grey Wolf and Whale Optimization algorithms to produce answers for scheduling issues in cloud tasks. A novel hybrid metaheuristic optimization approach is presented in [11] to address the coordination issue with directional overcurrent relays.

Among the recently popular swarm optimization algorithms, the Harris hawks optimization and the GWO have proven by many studies that they have the potential to solve many optimization problems and engineering problems [1, 4]. When these or other algorithms alone are not sufficient to solve certain optimization problems, hybridization of algorithms has been attempted.

In this study, a new optimization algorithm was developed by hybridizing GWO and HHO to obtain more successful optimization results, and this hybrid algorithm was named GWOHHO. F1-F13 test functions were used to measure GWOHHO performance. It has been seen that the GWOHHO algorithm provides better results than both GWO and HHO algorithms. In addition, these results were compared with other methods in the literature and it was seen that the GWOHHO algorithm generally obtained better results than the literature. In the rest of the article, the GWO and HHO methods are explained in Sections 2 and 3, respectively. In Section 4, the GWOHHO algorithm describes. In Section 5, the results of the GWOHHO method and their comparison with the literature are given. Finally, the conclusions of the GWOHHO method are given in Section 6.

GREY WOLF OPTIMIZATION (GWO)

Based on the hunting techniques developed by grey wolves in accordance with their social hierarchy, GWO optimization approach was created by Mirjalili et al. [4]. According to their hierarchical position, grey wolves made consist of four groups: alpha(α), beta(β), delta(δ), and omega(ω) as seen in Figure 1. The alpha group, which includes a female and a male wolf, is in charge of making crucial decisions like hunting. The second-ranked or "beta" wolves in the hierarchy assist the alpha wolves in carrying out their judgments. Delta wolves are the third-ranked wolf in the pack and follow the alpha and beta wolves' instructions. Since omegas are at the bottom of the wolf food chain, they are the last to be permitted to ingest.
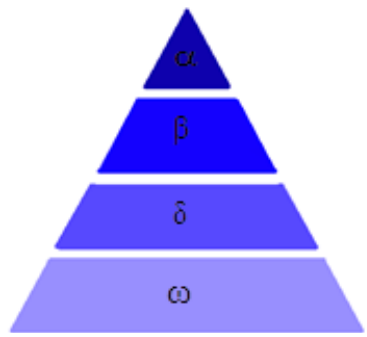
Figure 1. The representation of Grey wolves' hierarchy

The four stages of the GWO method include encircling the prey, hunting, attacking the prey (exploitation), and looking for prey (exploration).

Equations (1) and (2) describe the encirclement of prey by grey wolves.

$$\vec{H}(t+1)=\vec{H}l(t)-\vec{A}.\vec{D} \tag{1}$$

$$\vec{D}=\left|\vec{C}.\vec{H}l(t)-\vec{H}(t)\right| \tag{2}$$

The location of a grey wolf is $\vec{H}(t+1)$. $\vec{H}l(t)$ represents where the prey was at iteration i. From Equations (3) and (4), the coefficient vectors A and C are generated.

$$\vec{A}=2\vec{a}.\vec{r_1}+\vec{a} \tag{3}$$

$$\vec{C}=2.\vec{r_2} \tag{4}$$

$\vec{r_1}$ and $\vec{r_2}$ variables consist of random numbers between (0, 1). From 2 to 0, the a value decreases linearly. The closest new positions of the best agent can be checked by altering the A and C vectors.

The hunting phase follows the encircling phase and starts the process of finding the best answer. During this stage, alpha directs the hunt, with beta and delta occasionally joining in. Consequently, using Equations (5), (6), and (7) the three best positions are employed to update the position of the grey wolves.

$$\vec{D}_\alpha=\left|\vec{C_1}.\vec{H}_\alpha-\vec{H}\right|,\vec{D}_\beta=\left|\vec{C_2}.\vec{H}_\beta-\vec{H}\right|,\vec{D}_\delta=\left|\vec{C_3}.\vec{H}_\delta-\vec{H}\right| \tag{5}$$

$$\vec{H_1}=\vec{H}_\alpha-\vec{A_1}.\vec{D}_\alpha,\vec{H_2}=\vec{H}_\beta-\vec{A_2}.\vec{D}_\beta,\vec{H_3}=\vec{H}_\delta-\vec{A_3}.\vec{D}_\delta \tag{6}$$

$$\vec{H}(t+1)=\frac{\vec{H_1}+\vec{H_2}+\vec{H_3}}{3} \tag{7}$$

To sum up, the haphazard population is produced by the GWO algorithm. Delta, alpha, and beta wolves predict the location of the potential prey. The candidate

solution's distance is then updated. Then, to emphasize exploration and exploitation, respectively, *a* is decreased from 2 to 0. If *A<1*, they then go to the prey. They withdraw from attacking the prey if *A>1*. When GWO has reached a satisfactory conclusion, it is ended [4]. Figure 2 provides the GWO algorithm's pseudocode.

Initialize the grey wolf population H_i (i = 1, 2, ..., n)
Initialize a, A, C
Calculate the fitness of each search agent
H_α= The best search agent
H_β= The second best search agent
H_δ= The third best search agent
  while (t < Maximum iteration number)
    for each search agent
      Update the position of the current search agent using Eq (7)
    end for
  Update a, A, C
  Calculate the fitness of all search agent
  Update H_α, H_β, H_δ
  t=t+1
  end while
return H_α

Figure 2. Pseudocode of the GWO algorithm [4]

HARRIS HAWKS OPTIMIZATION (HHO)

The HHO algorithm is an optimization technique that mimics the action of multiple hawks working together to strike from various angles in order to confuse prey [1]. Hawks have developed a variety of hunting techniques in response to their prey's escape strategies. HHO is created as a result of the mathematical modeling of these distinct tactics. Given the right formulation, HHO can be used to solve any optimization issue because it uses an optimization method without gradients.

Diverse assault strategies, surprise pounce, and exploring prey of Harris hawks in [1] are used to simulate the phases of HHO.

Exploration phase

The Harris' hawks in HHO perch at random places and use one of two methods to wait for prey to appear using Equation (8).

$$H(t+1)=\begin{cases} H_{rand}(t)-r_1\left|H_{rand}(t)-2r_2H(t)\right| & q \geq 0.5 \\ (H_{rabbit}(t)-H_m(t))-r_3(LB+r_4(UB-LB)) & q < 0.5 \end{cases} \quad (8)$$

In each iteration, *r1, r2, r3, r4*, and *q* are updated and these variables are random numbers inside (0,1). *H(t + 1)* is hawks' position vector in the following *t*'s iteration. $H_{rabbit}(t)$ is the rabbit's position while *H(t)* is the hawks' present position vector. H_m is the current hawk population's mean position while $H_{rand}(t)$ is a hawk from the current population chosen at random. The variables' upper and lower bounds are

displayed in *LB* and *UB*. Using Equation (9), one can determine the hawks' typical position:

$$H_m(t) = \frac{1}{N} \sum_{i=1}^{N} H_i(t) \tag{9}$$

where hawks' total number is *N*. In addition, the whereabouts of each hawk in iteration *t* is represented by $H_i(t)$.

Exploration to Exploitation Transition

Using Equation (10), the rabbit's energy is represented as:

$$E = 2E_0\left(1 - \frac{t}{T}\right) \tag{10}$$

where *T* represents how many iterations are allowed. E_0 is its energy beginning state while *E* is the prey's escape energy.

Exploitation Phase

At this stage, there are 4 types of besiege.

Soft Besiege

Using Equations (11-12), the following set of rules can be used to describe soft besiege:

$$H(t+1) = \Delta H(t) - E\left|JH_{rabbit}(t) - H(t)\right| \tag{11}$$

$$\Delta H(t) = H_{rabbit}(t) - H(t) \tag{12}$$

where $J = 2(1 - r_5)$ represents the rabbit's randomized leap power when it is running away. In addition, r_5 represents a random number consisting of (0,1). *H(t)* stands for the difference between the present location and the position vector of the rabbit in cycle *t*. To mimic the nature of rabbit locomotion, the *J* value fluctuates at random during each repetition.

Hard Besiege

In hard besiege case, Equation (13) is used to update the current positions:

$$H(t+1) = H_{rabbit}(t) - E\left|\Delta H(t)\right| \tag{13}$$

Soft Besiege with Progressive Rapid Dives

The hawks are considered to be able to determine their next move based on the following rule in Equation (14) to launch a soft besiege.

$$Y = H_{rabbit}(t) - E\left|JH_{rabbit}(t) - H(t)\right| \tag{14}$$

According to the rule represented by Equation (15), they are expected to dive relying on the LF-based predictions.

$$Z = Y + S \times LF(D) \tag{15}$$

where S is a random vector of size $1xD$, D is the problem's dimension, and LF denotes the function of levy flight, which is derived using Equation (16):

$$LF(h) = 0.01 \times \frac{u \times \sigma}{|v|^{\frac{1}{\beta}}}, \quad \sigma = \left\{ \frac{\Gamma(1+\beta) \times \sin(\frac{\pi\beta}{2})}{\Gamma(\frac{1+\beta}{2}) \times \beta \times 2^{(\frac{\beta-1}{2})}} \right\}^{\frac{1}{\beta}} \tag{16}$$

where u, v are random values within of $(0, 1)$, and where β (default constant) is set to 1.5. As a result, Equation (17) can be used to update the hawks' positions during the soft besiege phase.

$$H(t+1) = \begin{cases} Y & \text{if } F(Y) < F(H(t)) \\ Z & \text{if } F(Z) < F(H(t)) \end{cases} \tag{17}$$

Here, Y and Z are determined using Equations (14-15).

Hard Besiege with Progressive Rapid Dives
In Equation (18), the following rule is applied.

$$H(t+1) = \begin{cases} Y & \text{if } F(Y) < F(H(t)) \\ Z & \text{if } F(Z) < F(H(t)) \end{cases} \tag{18}$$

where the new rules in Equations (19) and (20) determine Y and Z. The HHO's pseudocode is given in Figure 3.

$$Y = H_{rabbit}(t) - E \left| JH_{rabbit}(t) - H_m(t) \right| \tag{19}$$

$$Z = Y + S \times LF(D) \tag{20}$$

Inputs: The population size N and maximum number of iterations T
Outputs: The location of rabbit and its fitness value

Initialize the random population $H_i (i = 1, 2, . . . , N)$

while (stopping condition is not met) **do**
   Calculate the fitness values of hawks
  Set H_{rabbit} as the location of rabbit (best location)
  for (each hawk (H_i)) **do**
   Update the initial energy E_0 and jump strength J
   Update the E using Eq. (10)

   if ($|E| \geq 1$) **then**
    Update the location vector using Eq. (8)
   if ($|E| < 1$) **then**

    if ($r \geq 0.5$ and $|E| \geq 0.5$) **then**
     Update the location vector using Eq. (11)
    else if ($r \geq 0.5$ and $|E| < 0.5$) **then**
     Update the location vector using Eq. (13)
    else if ($r < 0.5$ and $|E| \geq 0.5$) **then**
     Update the location vector using Eq. (17)
    else if ($r < 0.5$ and $|E| < 0.5$) **then**
     Update the location vector using Eq. (18)

Return H_{rabbit}

Figure 3. HHO's pseudocode [1]

THE PROPOSED METHOD

GWO and HHO algorithms may not be sufficient to reach the target value in the solutions of benchmark functions and engineering problems. With the algorithm we developed, a hybrid algorithm called GWOHHO, which highlights the good features of GWO and HHO, is proposed. The pseudocode of the GWOHHO algorithm is given in Figure 4.

In the proposed GWOHHO method, fitness functions for both HHO and GWO are calculated after the initial population is created. And then GWO and HHO algorithms run their own processing steps. In both algorithms, new fitness function value, best position, new population values and positions are obtained. In the next iteration, the population values of the algorithm with the best fitness value are given as the input population to both algorithms and these steps are repeated until the specified maximum iteration is reached.

```
Create random initialize population X_k(k = 1, 2, . . . , N)
r=0
while (r<T) do
    [Alpha_score, Alpha_pos, gPositions]=GWO(Positions, fobj, r);
    [Rabbit_Energy, Rabbit_Energy_pos, hPositions]=HHO(Positions, fobj, r);
    if Alpha_score < Rabbit_Energy
    Best_score = Alpha_score;
    Best_pos = Alpha_pos;
    Positions=gPositions;
    else
    Best_score = Rabbit_Energy;
    Best_pos = Rabbit_Energy_pos;
    Positions=hPositions;
    end if
    r=r+1
end while
Return Best_score, Best_pos
```

Figure 4. Pseudocode of the proposed GWOHHO

RESULTS AND DISCUSSION

The GWOHHO method was tested with the F1-13 benchmark functions which are composed of unimodal and multimodal. These functions are used to compare the proposed algorithm with different methods in the literature. The same parameters were used as much as possible in order to make an accurate comparison with the literature and so it has been tried to reach the articles containing the same parameters. In the literature, the population number, the maximum iteration number, and independent run number for GWO, HHO, WOA and PSO algorithms were used as 30, 500, and 30, respectively [1, 4, 5]. In order to compare with the literature, same were taken parameters (the population number=30, the maximum iteration number=500, and independent run number=30). Mean and standard deviation results were obtained from these run results. GWO, HHO, WOA and PSO methods were compared with the proposed GWOHHO and it was generally found that better results were obtained. Unimodal F1-F7 functions are given in Table 1. *Lb* is a low boundary while *Ub* is an upper boundary. *Dim* refers to the function size. In the exploitation process, F1-F7 functions can measure the performance of the optimization algorithm and its capability.

Table 1. Unimodal F1-F7 functions

Func. Num.	Function	Range		Dim
		Lb	Ub	
$F_1(b)$	$\sum_{i=1}^{n} b_i^2$	-100	100	30
$F_2(b)$	$\sum_{i=1}^{n}|b_i| + \prod_{i=1}^{n}|b_i|$	-10	10	30

$$F_3(b) \quad \sum_{i=1}^{n}(\sum_{j-1}^{i} b_j)^2 \qquad -100 \quad 100 \quad 30$$

$$F_4(b) \quad \max_i \{|b_i|, 1 \le i \le n\} \qquad -100 \quad 100 \quad 30$$

$$F_5(b) \quad \sum_{i=1}^{n-1}[100(b_{i+1} - b_i^2)^2 + (b_i - 1)^2] \qquad -30 \quad 30 \quad 30$$

$$F_6(b) \quad \sum_{i=1}^{n}([b_i + 0.5])^2 \qquad -100 \quad 100 \quad 30$$

$$F_7(b) \quad \sum_{i=1}^{n} i b_i^4 + random[0,1) \qquad -1.28 \quad 1.28 \quad 30$$

The results of the F1-F7 functions are shown in Table 2. Here, σ stands for the standard deviation whereas μ indicates the average value. The GWOHHO algorithm obtained the optimal result for 7 out of 7 functions for unimodal comparison functions. In light of the results obtained, it is seen that the developed hybrid GWOHHO algorithm is successful for unimodal comparison functions.

Table 2. Results of F1-F7 functions

F	PSO[5]		HHO[1]		GWO[4]		WOA[5]		GWOHHO	
	μ(ave)	σ(std)	μ(ave)	σ(std)	μ(ave)	σ(std)	μ(ave)	σ(std)	μ(ave)	σ(std)
F1	1.3600E-04	2.0200E-04	3.95E−97	1.72E−96	6.5900E-28	6.3400E-05	1.4100E-30	4.9100E-30	**5.6430E-103**	2.1046E-102
F2	4.2144E-02	4.5421E-02	1.56E−51	6.98E−51	7.1800E-17	2.9014E-02	1.0600E-21	2.3900E-21	**3.5500E-52**	6.1812E-52
F3	7.0126E+01	2.2119E+01	1.92E−63	1.05E−62	3.2900E-06	7.9150E+01	5.3900E-07	2.9300E-06	**6.2382E-66**	3.2598E-65
F4	1.0865E+00	3.1704E-01	1.02E−47	5.01E−47	5.6100E-07	1.3151E+00	7.2581E-02	3.9747E-01	**2.4655E-53**	1.3235E-52
F5	9.6718E+01	6.0116E+01	1.32E−02	1.87E−02	2.6813E+01	6.9905E+01	2.7866E+01	7.6363E-01	**2.0000E-03**	4.5000E-03
F6	1.0200E-04	8.2800E-05	1.15E−04	1.56E−04	8.1658E-01	1.2600E-04	3.1163E+00	5.3243E-01	**8.1795E-06**	1.7532E-05
F7	1.2285E-01	4.4957E-02	1.40E−04	1.07E−04	2.2130E-03	1.0029E-01	1.4250E-03	1.1490E-03	**2.2053E-05**	2.1903E-05

The best search and objective space representations of F1-F7 functions are given in Figure 5. In F1-F7 functions, before reaching the maximum iteration value, it gets stuck at a value close to the target value, or convergence decreases.
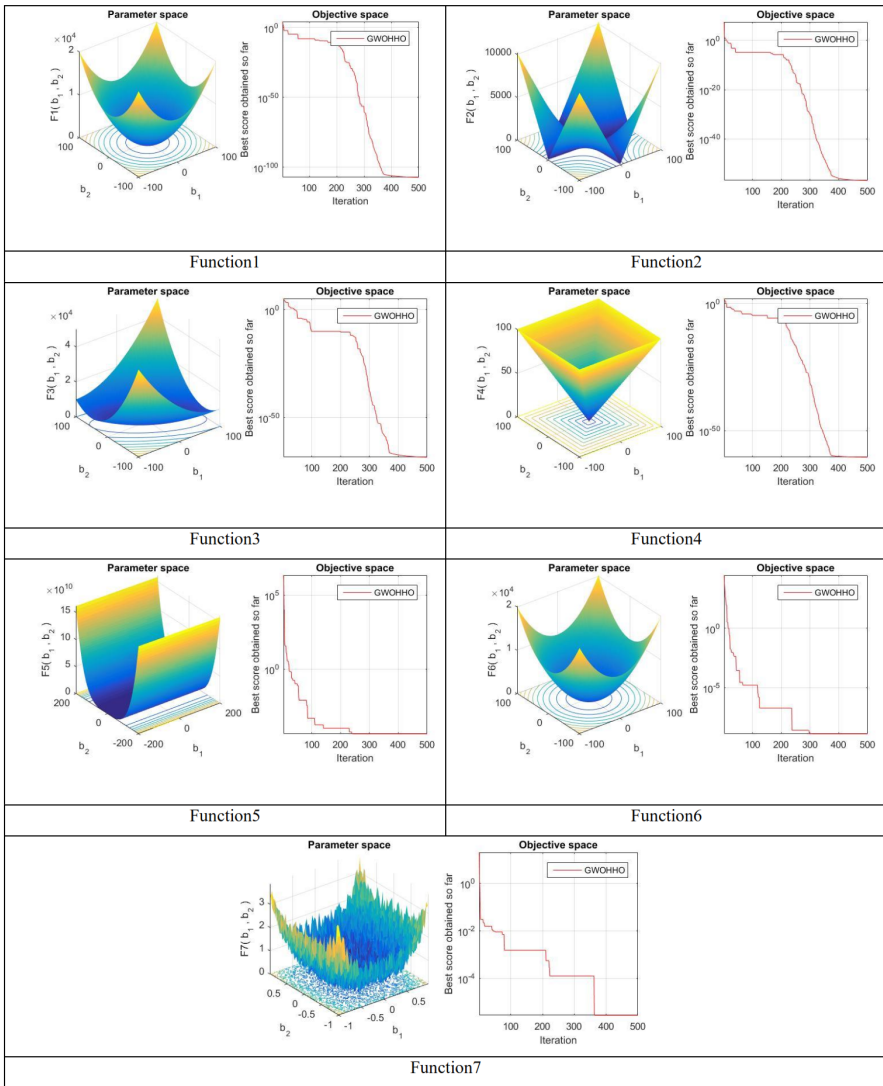
Figure 5. The best search and objective space representations of F1-F7 functions

Multimodal F8-F13 functions are given in Table 3. Since an optimization algorithm must avoid all local optima to reach the global optimum, F8-F13 functions can measure the strength of an optimization algorithm in terms of avoiding local optima.

Table 3. Multimodal F8-F13 functions

Func. Num.	Function	Range		Dim
		Lb	Ub	
$F_8(b)$	$\sum_{i=1}^{n} -b_i \sin(\sqrt{|b_i|})$	-500	500	30
$F_9(b)$	$\sum_{i=1}^{n} \left[b_i^2 - 10\cos(2\pi b_i + 10) \right]$	-5.12	5.12	30
$F_{10}(b)$	$-20\exp(-0.2\sqrt{\frac{1}{n}\sum_{i=1}^{n} b_i^2}) - \exp(\frac{1}{n}\sum_{i=1}^{n}\cos(2\pi b_i)) + 20 + e$	-32	32	30
$F_{11}(b)$	$\frac{1}{4000}\sum_{i=1}^{n} b_i^2 - \prod_{i=1}^{n}\cos(\frac{b_i}{\sqrt{i}}) + 1$	-600	600	30
$F_{12}(b)$	$\frac{\pi}{n}\left\{ 10\sin(\pi y_1) + \sum_{i=1}^{n-1}(y_i - 1)^2\left[1 + 10\sin^2(\pi y_{i+1})\right] + (y_n - 1)^2 \right\}$ $+\sum_{i=1}^{n} u(b_i, 10, 100, 4)$ $y_i = 1 + \frac{b_i + 1}{4}$ $u(b_i, a, k, m) = \begin{cases} k(b_i - a)^m & b_i > a \\ 0 & -a < b_i < a \\ k(-b_i - a)^m & b_i < -a \end{cases}$	-50	50	30
$F_{13}(b)$	$0.1\left\{ \sin^2(3\pi b_1) + \sum_{i=1}^{n}(b_i - 1)^2\left[1 + \sin^2(3\pi b_i + 1)\right] \right.$ $\left. + (b_n - 1)^2\left[1 + \sin^2(2\pi b_n)\right] \right\} + \sum_{i=1}^{n} u(b_i, 5, 100, 4)$	-50	50	30

The results of F8-F13 functions are given in Table 4. The GWOHHO algorithm obtained the optimal result for 5 out of 6 functions for multimodal comparison functions. According to the F8-F13 results, more than half of the results have been improved. F10 achieved the best result in HHO. Other functions achieved the best result in GWOHHO.

Table 4. Results of F8-F13 functions

F	PSO[5]		HHO[1]		GWO[4]		WOA[5]		GWOHHO	
	μ(ave)	σ(std)	μ(ave)	σ(std)	μ(ave)	σ(std)	μ(ave)	σ(std)	μ(ave)	σ(std)
F8	-4.8413E +03	1.1528E +03	-1.25E +04	1.47E +02	-6.1231E +03	-4.0874E +03	-5.0808E +03	6.9580E +02	**-1.2569E +04**	1.0091E +00
F9	4.6704E +01	1.1629E +01	**0.00E +00**	0.00E +00	3.1052E -01	4.7356E +01	**0.0000E +00**	0.0000E +00	**0.0000E +00**	0.0000E +00
F10	2.7602E -01	5.0901E -01	**8.88E −16**	4.01E −31	1.0600E -13	7.7835E -02	7.4043E +00	9.8976E +00	8.8818E -16	0.0000E +00
F11	9.2150E -03	7.7240E -03	**0.00E +00**	0.00E +00	4.4850E -03	6.6590E -03	2.8900E -04	1.5860E -03	**0.0000E +00**	0.0000E +00
F12	6.9170E -03	2.6301E -02	2.08E −06	1.19E −05	5.3438E -02	2.0734E -02	3.3968E -01	2.1486E -01	**9.2137E −07**	1.6442E -06
F13	6.6750E -03	8.9070E -03	1.57E −04	2.15E −04	6.5446E -01	4.4740E -03	1.8890E +00	2.6609E -01	**1.8052E −05**	5.7112E -05

The best search and objective space representations of F8-F13 functions are given in Figure 6. Convergence to the target value occurs before reaching the maximum iteration value in F9 and F11 functions. In other functions, before reaching the maximum iteration value, it gets stuck at a value close to the target value, or convergence decreases.

CONCLUSION

A hybrid algorithm named GWOHHO is proposed that highlights the good features of GWO and HHO. In this way, the results of the benchmark test functions have been improved, that is, it has been seen that it is more successful in reaching the target value. In the end, a more efficient hybrid optimization algorithm is obtained. To test the effectiveness of this algorithm, 13 mathematical test functions, including unimodal and multimodal functions, were used. The average fitness and standard deviation values were calculated as a result of running the proposed method 30 times. Search and objective space representations of the best of these 30 runs are also given.

In addition, the mean value obtained from the proposed GWOHHO was contrasted with the GWO, WOA, HHO and PSO methods in the literature. According to the algorithms compared with the literature, GWOHHO was found to achieve the best results in 7 of 7 tests for unimodal benchmark functions and 5 out of 6 tests for multimodal benchmark functions. As a result, the proposed algorithm generally outperforms the literature results.
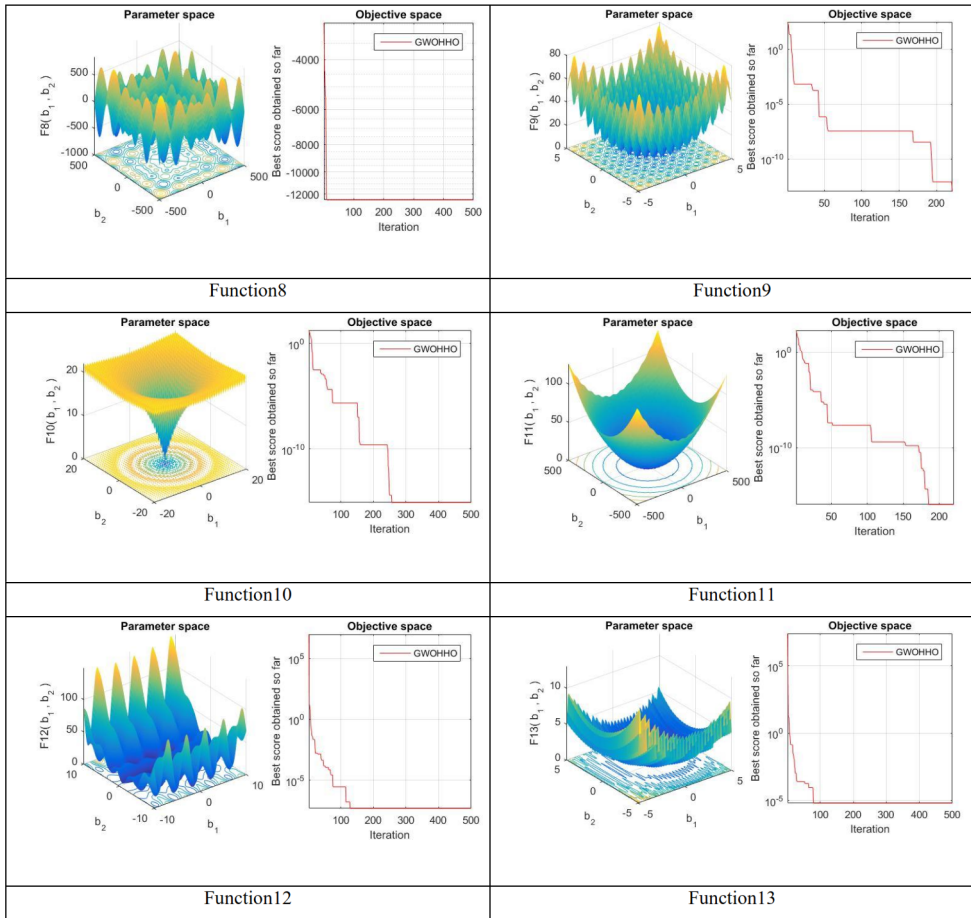
Figure 6. The best of search and objective space representations of F8-F13 functions

According to the results, it is seen that the GWOHHO method is promising and can be applied in engineering applications. In the future, we aim to further develop different hybrid varieties of the proposed method and use it in engineering applications.

ACKNOWLEDGMENTS

REFERENCES

[1] A. A. Heidari, S. Mirjalili, H. Faris, I. Aljarah, M. Mafarja, and H. L. Chen, "Harris hawks optimization: Algorithm and applications," *Future Generation Computer Systems-the International Journal of Escience,* vol. 97, pp. 849-872, Aug 2019, doi: 10.1016/j.future.2019.02.028.

[2] R. Abbassi, A. Abbassi, A. A. Heidari, and S. Mirjalili, "An efficient salp swarm-inspired algorithm for parameters identification of photovoltaic cell models," *Energy Conversion and Management,* vol. 179, pp. 362-372, Jan 2019, doi: 10.1016/j.enconman.2018.10.069.

[3] H. Faris *et al.*, "An intelligent system for spam detection and identification of the most relevant features based on evolutionary Random Weight Networks," *Information Fusion,* vol. 48, pp. 67-83, Aug 2019, doi: 10.1016/j.inffus.2018.08.002.

[4] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey Wolf Optimizer," (in English), *Advances in Engineering Software,* vol. 69, pp. 46-61, Mar 2014. [Online]. Available: <Go to ISI>:// WOS:000331924800006.

[5] S. Mirjalili and A. Lewis, "The Whale Optimization Algorithm," (in English), *Advances in Engineering Software,* vol. 95, pp. 51-67, May 2016. [Online]. Available: <Go to ISI>:// WOS:000371899900006.

[6] S. Arora and P. Anand, "Binary butterfly optimization approaches for feature selection," (in English), *Expert Systems with Applications,* vol. 116, pp. 147-160, Feb 2019. [Online]. Available: <Go to ISI>://WOS:000449240500012.

[7] D. Karaboga and B. Basturk, "On the performance of artificial bee colony (ABC) algorithm," (in English), *Applied Soft Computing,* vol. 8, no. 1, pp. 687-697, Jan 2008. [Online]. Available: <Go to ISI>://WOS:000249508500060.

[8]   J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of ICNN'95-international conference on neural networks*, 1995, vol. 4: IEEE, pp. 1942-1948.

[9] G. Y. Zhu and W. B. Zhang, "Optimal foraging algorithm for global optimization," (in English), *Applied Soft Computing,* vol. 51, pp. 294-313, Feb 2017. [Online]. Available: <Go to ISI>:// WOS:000395836900022.

[10] J. Ababneh, "A Hybrid Approach Based on Grey Wolf and Whale Optimization Algorithms for Solving Cloud Task Scheduling Problem," (in English), *Mathematical Problems in Engineering,* vol. 2021, Sep 13 2021. [Online]. Available: <Go to ISI>://WOS:000701217500008.

[11] A. Korashy, S. Kamel, F. Jurado, and A. R. Youssef, "Hybrid Whale Optimization Algorithm and Grey Wolf Optimizer Algorithm for Optimal Coordination of Direction Overcurrent Relays," (in English), *Electric Power Components and Systems,* vol. 47, no. 6-7, pp. 644-658, Apr 21 2019. [Online]. Available: <Go to ISI>://WOS:000471437300001.

DEEP FOREST APPROACH FOR ZERO-DAY ATTACKS DETECTION

Mahmut TOKMAK [1]

INTRODUCTION

Today, computer technologies have entered every aspect of our lives. These technologies have become a part of human life with the concepts of developing internet networks, mobile technologies and internet of things (IoT). It also shapes people's lifestyles. Computer and internet are used in many areas such as daily life, education, banking, health, transportation, security, banking, shopping, communication, engineering and architecture. However, it can bring challenges as well as opportunities. These developments in the world of technology have revealed the fact that cyber attackers and cyber attacks are emerging and their number is increasing day by day.

In recent years, the internet has become a cyber world that contains threats to individuals and various organizations. Sophisticated applications and new technologies are developed every day to meet the changing needs of the business world and make life easier. However, malicious sources try various methods to exploit vulnerabilities in systems.[1,2]. Errors in a application and hardware components that allow an attacker to execute malicious commands, gain unauthorized access to data and/or perform various denial-of-service attacks are defined as vulnerabilities. [3]. Using different tools and methods, attackers can perform actions called. Attacks are actions taken from one or more computers to damage other computers or networks and disrupt their normal functioning, using various methods. Attackers may initiate these actions reach to their malicious ends, for personal gratification, for financial gain or reward [4,5]. Of these attacks, Malware is an abbreviation for "malicious software". It refers to malware developed by attackers to steal data, damage or destroy computers and computer systems.

Some of them are: virus, worm, spyware, adware, Trojan horse, botnet, rootkit, backdoor, Ransomware. Ransomware is a form of malware that prevents users from accessing their systems, personal files or devices by encrypting files on the system. [6,7]. Denial of Service (DoS/DDos) attacks: aims to consume the resources of the target system by using a large number of clients and to prevent the service provided by the target system [8]. Initially, attacks were made from a single source, but now attacks are carried out on target systems from multiple sources at the same time [9].

1   Mahmut TOKMAK, Bucak Zeliha Tolunay School of Applied Technology and Management, Department of Management Information Systems, Burdur, Orcid Id: 0000-0003-0632-4308, mahmuttokmak@mehmetakif.edu.tr

Phishing is one of the old and effective electronic fraud methods designed to steal users' personal information, credit card information, social media information and passwords on web applications. Phishing techniques use various communication methods such as SMS, MMS, e-mail, fake web pages [10]. SQL injection attack is a type of cyber attack carried out by exploiting vulnerabilities in systems with databases through SQL queries. SQL injection is an attack method for attacking database applications. Attacker adds new SQL statements to the input-related field in the standard application. This attack is an unauthorized access to the database and harm in a variety of ways, including manipulation and disclosure of sensitive information, data. SQL injection attacks can cause attackers to log on to the system, tamper with data, cancel or modify some operations, expose all data in the database, delete all data in the database. [5,11,12]. Cryptojacking is a form of attack that uses any person's mobile device or computer for the attacker's cryptocurrency mining [13,14]. Session hijacking and Man-in-the-middle (MITM) is a form of attack in which the attacker infiltrates the network of the attacked user and source, reading and exchanging messages. The critical situation in this attack is that the user is unaware of the man in the middle. The user thinks that he is directly connected to the network, but the network traffic passes through the man in the middle. In this way, the man in the middle can access and even change all the network data, login information and many personal data of the user. Zero-Day describes a cyber-attack technique that hackers use to attack systems by targeting vulnerabilities. The term zero-day is used here because the developer has just learned about the vulnerability. At this stage, hackers can easily exploit vulnerabilities knowing that there is no effective defense. This makes zero-day vulnerabilities a serious security threat [15].

Attackers update themselves and the software they use every day and create new attack scenarios. This means that both the types and numbers of attacks on data security are increasing [16]. Zero-day attacks start on the day the vulnerability of the software used is detected. When software developers detect such a vulnerability, they develop security patches to fix this vulnerability and ask users to update with these security patches. Attackers try to detect possible "Zero-Day" vulnerabilities in software and use this vulnerability to infiltrate victim systems. The time between the first exploitation of a vulnerability in a computer system or software product and when software developers and security professionals begin developing a response and blocking it is known as the vulnerability window [2].

With the increase in attacks on data security, the need for security technologies is constantly increasing. However, there is a rapid development in Intrusion Detection Systems (IDS).

The best way to prevent updated attacks is to update and improve IDS [16]. At the center of tackling the exponential increase in cyberattacks is IDS, which can detect zero-day cyber-attacks. The methods used in IDS can be examined under

two headings; Signature-Based Intrusion Detection, Anomaly-Based Intrusion Detection. Signature-based IDS, with distinctive features of previously occurring attacks; store information such as system calls, permissions, network access, and runtime in a signature database. They then classify programs or network streams using these pre-recorded signatures. In IDS that perform static analysis in this way, keeping the signature database up-to-date is critical for the success of the system. In addition, these systems fail against zero-day attacks even if they have an up-to-date database. Therefore, studies on Machine Learning (ML) based anomaly detection systems have gained importance in order to develop an intelligent IDS that can adapt to the changing nature of attacks and give successful results against new attacks [15,17].

Applications of ML techniques are used in various areas of life like education, medicine, finance, transport, manufacturing industry, automotive [18,19]. One of these areas is IDS. ML techniques have been widely used to design and build robust IDS.

Also, the high false positive detection rate of current available IDS is a problem, so its detection performance and real-life uses are limited. Consequently, many zero-day attacks remain identify and their damage increases.[20]. In order to develop IDS, new studies are constantly carried out in the literature and new methods are suggested [16].

Al-rimy et al. [21] proposed the Decision Fusion-Based Model to detect zero-day attacks. Two types of detection estimators are created in this model. The first type is a collection of ensemble behaviorally based classifiers, while the second type is an anomaly-based predictor. The decisions of both estimator methods were merged using the fusion method and established an early detection mechanism. Sharma et al. [22] proposed a context graph-based method to predict zero-day attacks. In their proposed method, they aimed to detect attacks on IoT networks as a target. They stated that the distributed approach is capable of efficiently mitigating zero-day threats, with 33% and 21% improvements compared to the centralized diagnostic system. Sun et al. [23] proposed a Bayesian network probability method to detect zero-day attack chains. They stated that in the attacks carried out by the attackers, a series of attacks occurred in order to reach the target. In their proposed method they used system calls and created a prototype using graph. Q. Zhou and Pezaros [24] tested and measured the performance of six ML methods. The authors used the CIC-AWS-2018 attack dataset in their tests. They used Random Forest (RF), Decision Tree (DT), K-Nearest Neighbor (KNN), Multilayer Perceptron (MLP), Quadratic Discriminant Analysis (QDA) and Naive Bayes (NB) classifiers. Zhao et al. [25], presented transfer learning to detect zero-day attacks using knowledge of known attacks to resolve hard-to-detect attacks in real time. They used feature-based approach linear transformation. In addition, the authors proposed a cluster

improved transfer learning approach to make it stronger in detecting zero-day attacks. Obeidat et al. [26], measured the performance of the model with various classifiers on the KDD'99 data set. Weka version 3.7.12 is used as software. For the test, 60.000 independent samples from the training dataset were taken. In the study; J48, RF, Random Tree (RT), Decision Table (DT), MLP, NB and Bayes Network classification algorithms are used. Blaise et al. [27] proposed an unsupervised port based method to detect botnets and zero-day attacks. Sameera and Shashi [28], used deep transudative transfer learning to detect unknown attacks. The authors performed various experiments using the NSL-KDD and CIDD datasets to measure the performance of the proposed model. Hindy vd [20] used ML and Deep Learning (DL) to detect vulnerabilities. They used the CICIDS2017 and NSL-KDD dataset in their study. Nkongolo et al [29] presented a novel dataset ,named UGRansome, for anomaly prediction containing zero-day attacks behaviours. Nkongolo et al [30] stated that single classifiers like SVM and NB are unstable for Zero-Day Threats. Therefore, they proposed a model using Ensemble Learning with a Genetic Algorithm (GA) optimizer with three ML algorithms NB, RF and SVM. Nkongolo et al [31] performed malware detection using the UGRansome dataset and the Boruta and PCA algorithms..

Many performance evaluation datasets such as KDDCUP'99, NSL-KDD, CICIDS2017, CIC-AWS-2018, UGRansome are used to detect Zero-Day attacks. One of these datasets, KDDCUP'99, is the dataset commonly used by researchers working on network anomaly detection. However, it was determined that some errors found in the KDDCUP'99 dataset affected the evaluations made on this dataset, and then the NSL-KDD dataset, which is the error-free form of the KDDCUP'99 dataset, was created. The NSL-KDD dataset is actively used by researchers working on attack detection and anomaly detection [32]. The UGRansome dataset is a dataset that has been used since 2021. Therefore, studies published using the UGRansome dataset are few in the literature [29,29,30].

In this study, a newly method, the Deep Forest algorithm, and a newly dataset, UGRansome, were used. In order to compare the performance of the proposed method, experiments were conducted with the NSL-KDD dataset, which is more studied than the UGRansome dataset, and the results are presented.

MATERIAL AND METHOD

Dataset

Ransomware is called new malware or zero-day attacks. Ransomware can target different platforms, particularly networks. Ransomware can target different platforms, particularly computer networks. Therefore, it is important to detect these malware type which have a wide range of malicious effects. UGRansome dataset was used in this study. UGRansome was created in 2021 and is publicly accessible. UGRansome consists of different families of ransomware. UGRansome dataset was

created using the UGR'16 and ransomware datasets. It consists of 14 bundles of attributes and 207.534 features [29].
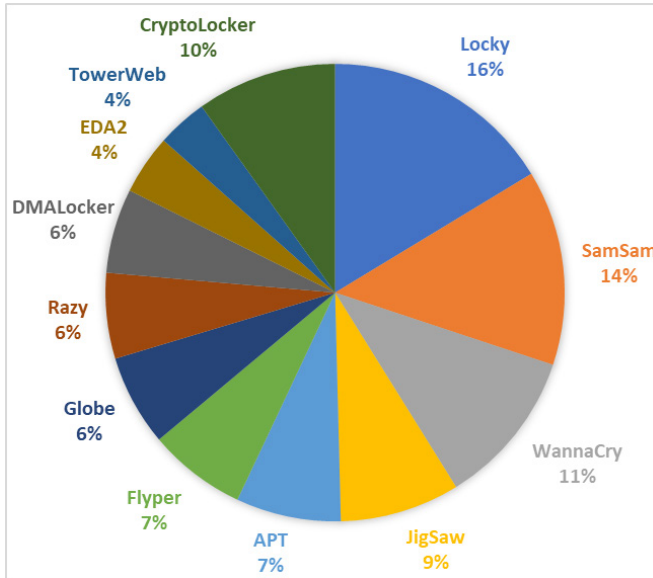
Figure 1. Attack Classes of the UGRansome Dataset

The UGRansome dataset has 12 different Ransomeware families and these classes are shown in Figure 1.

Table 1. UGRansome Dataset Structure

Attribute	Example	Description
Prediction	A	Anomaly
Ransomware	Locky	Novel malware
Bitcoins (BTC)	1964 BTC	Ransome payment
Dollars (USD)	520 $	Ransome payment
Cluster	8	Group assigned
Seed address	1DA11mPS	Malware address
Expended address	1BonuSr7	Malware address
Port	5061	Communication port
Malware	DoS	Novel malware
Network traffic	6103 bytes	Periodic network flow
IP address	Class B	Unique address
Flag	AF	Network state
Protocol	UDP	Communication rule
Timestamp	12 s	Traffic duration

Data structure of the UGRansome Zero-Day dataset is shown in Table 1. These features are triangulated using data fusion and are well suited for anomaly detection. The various malware are divided into three prediction classes: signature (S), synthetic signature (SS) and anomaly (A). The numbers of prediction classes are shown in the Figure 2 [31].
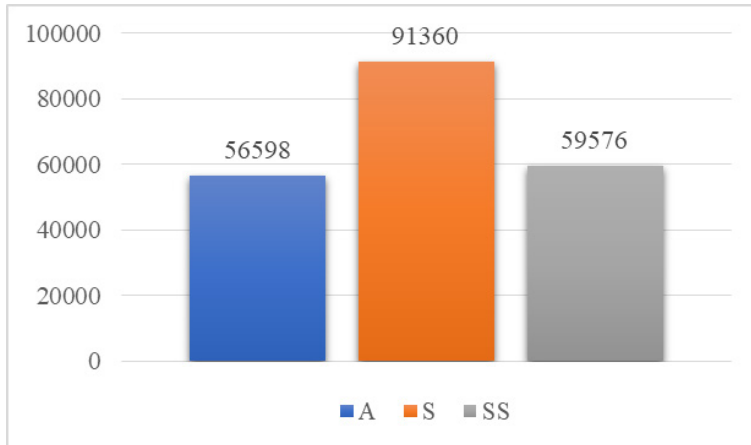
Figure 2. Prediction Classes

In order to compare the performances of the model created in this study, the NSL-KDD dataset, which is the derived version of the KDDCUP'99 dataset, was also used. It has been determined that the KDDCUP'99 dataset, which is used to compare the success of intrusion detection methods, contains too much repetitive data, affects the success rates of the algorithms and causes much higher results than normal [33]. As a result, the NSL-KDD dataset was created by clearing the errors from the KDDCUP'99 dataset. Each data in the NSL-KDD dataset has 41 attributes. Three features are categorical and 38 features are numerical.
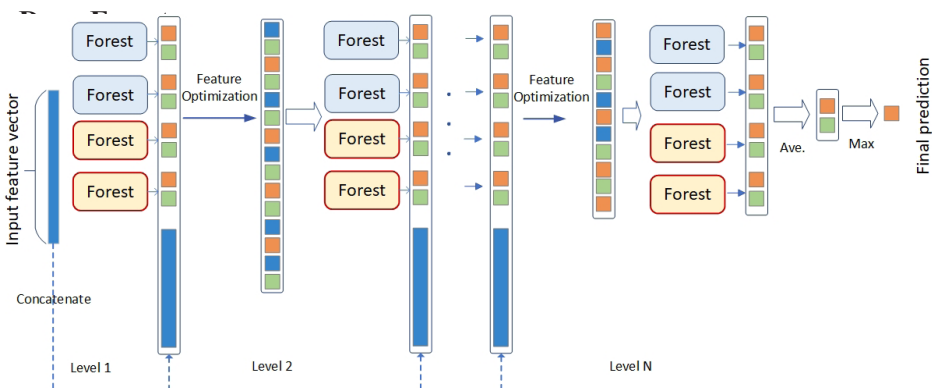
Figure 3. Architecture of Deep Forest

Deep Forest; layered architecture was suggested by Zhou and Feng. Representative learning in DNN layers, the features related to the problem to be

solved are learned and these learned features form the inputs of the following layer. Inspired by this learning, Deep Forest uses a cascade structure, where each layer receives the feature information processed by the previous layer and delivers the machining result to the next layer. Here different types of forests are included to ensure diversity. Because diversity is essential for community building [34]. In a Deep Forest structure as shown in Figure 3, each level of the cascade consists of two RF (grey) and two completely RF (red). Suppose there are two classes to guess; each forest yields a two-dimensional vector of classes, which are then combined to represent the original input [35].

EXPERIMENTS AND RESULTS

The design of the models established for classification with the Deep Forest algorithm was carried out using the Python programming language. In experiments on the model, 80% of the data set is reserved for training and 20% for testing. Python numpy, sklearn, pandas, statistics, matplotlib.pyplot libraries were used in processes such as preliminary file uploading for classification and data visualization in the calculation of performance metrics. Deep Forest Python library is used for the proposed Deep Forest architecture [36].

Used to evaluate the performance of the established model; the accuracy value, which is the ratio of correctly classified class samples, is in Equation 1, the precision value, which expresses how many of the positively predicted samples are actually positive, is in Equation 2, the sensitivity (recall) value, which expresses how much of the true positive values are correct, is shown in Equation 3, and the F1-score, which is the harmonic mean of precision and sensitivity values, is shown in Equation 4.

$$Accuracy = \frac{TP + TN}{TP + FN + TN + FP} \tag{1}$$

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{3}$$

$$F1 - Score = \frac{2xSensitivityxPrecision}{Sensitivity + Precision} \tag{4}$$

The performance values obtained as a result of the experiments with the Deep Forest and UGRansome dataset are presented in Table 2. As seen in Table 2, 97.7% of high performance criteria were obtained. In addition, the complexity matrix is shared in Figure 4.

Table 2. Deep Forest Performance Metrics with UGRansome

	Precision	Recall	F1-score	Support
A	0.996	0.996	0.996	11343
S	0.996	0.997	0.996	18223
SS	0.998	0.998	0.998	11941
Accuracy			0.997	41507
Macro avg	0.997	0.997	0.997	41507
Weighted avg	0.997	0.997	0.997	41507
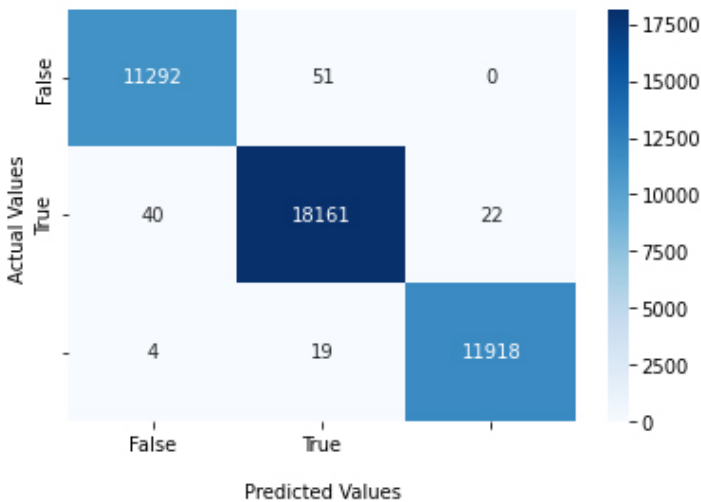
Figure 4. Confusion Matrix for Deep Forest and UGRansome

To demonstrate the performance of the Deep Forest technic, experiments were conducted with the current ML methods and the UGRansome dataset. In these experiments, Support VectorMachine (SVM), Adaboost, NB, RF, DNN algorithms were used. In the test results in the Table 3, 96.41%, 85.28%, 70.08%, 96.09% and 97.47% were obtained, respectively.

Table 3. ML Algorithms Performance Metrics with UGRansome

ML Algorithms	Accuracy	Precision	F1 Score	Recall
SVM	0.964102	0.961986	0.963144	0.964367
Adaboost	0.852892	0.844032	0.842606	0.841359
NB	0.700894	0.681798	0.669276	0.681005
RF	0.960946	0.962266	0.958832	0.956975
DNN	0.974607	0.974083	0.974145	0.974234

The results of the experiments applied to test the performance of the Deep Forest architecture on a different data set are shown in the Table 4. The proposed algorithm also achieved 99.9% results on the NSL-KDD dataset and obtained a strong accuracy. In addition, the complexity matrix is shared in Figure 5.

Table 4. Deep Forest Performance Metrics with NSL-KDD

	Precision	Recall	F1-score	Support
Anomaly	0.999	0.999	0.999	11651
Normal	0.999	0.999	0.999	13544
Accuracy			0.999	25195
Macro avg	0.999	0.999	0.999	25195
Weighted avg	0.999	0.999	0.999	25195
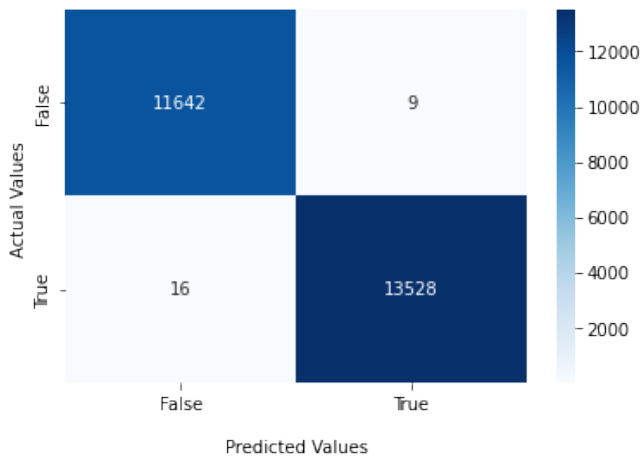
Figure 5. Confusion Matrix for Deep Forest and NSL-KDD

Experiments were made with the existing ML methods of the Deep Forest algorithm and the NSL-KDD data set. SVM Adaboost, NB, RF, DNN algorithms were used in these experiments. In the test results in the table, 99.49%, 98.49%, 87.01%, 97.61% and 99.69% were obtained, respectively.

Table 4. ML Algorithms Performance Metrics with NSL-KDD

ML Algorithms	Accuracy	Precision	F1 Score	Recall
SVM	0.994959	0.994948	0.994934	0.994919
Adaboost	0.984997	0.984931	0.984920	0.984909
NB	0.870173	0.878223	0.870133	0.875857
RF	0.976106	0.978460	0.975894	0.974361
DNN	0.996944	0.996865	0.996929	0.996994

CONCLUSION

Technology and the internet have become an indispensable element of both individuals and institutions and organizations. The widespread use of the internet and technology brings with it many security vulnerabilities. These vulnerabilities make networks and systems vulnerable to hackers, viruses, and other attacks. IDS systems developed to prevent Zero-Day attacks are one of the important areas of practice today.

In this study, Deep Forest approach with a layered architecture was used for the detection of Zero-Day attacks. In order to measure and compare the performance of the method, the UGRansome dataset, which can be considered new, and the NSL-KDD dataset, which is widely used in the literature, were used. First, experimental studies were performed with the UGRansome dataset. In experiments with the Deep Forest algorithm and the UGRansome dataset, an accuracy value of 97.7% was obtained. The performances of the same dataset and existing ML algorithms were evaluated. It was seen that the Deep Forest algorithm achieved more successful results on the UGRansome dataset than the SVM, Adaboost, NB, DNN methods.

Then, experimental studies were performed with the NSL-KDD dataset. In experiments with Deep Forest algorithm and NSL-KDD dataset, 99.9% accuracy value was obtained. The performances of the same dataset and existing ML algorithms were evaluated. It was seen that the Deep Forest algorithm achieved more successful results on the NSL-KDD dataset than the SVM, Adaboost, NB, DNN methods.

As a result of the experiments, it was concluded that systems based on Deep Forest methods can be used as a successful decision support system in the detection of Zero-Day attacks. In future studies, it is aimed to conduct experimental studies with different ML methods and different data sets.

REFERENCES

[1] Ahmetoğlu, H., & Daş, R. (2021). Makine Öğrenmesi Yöntemleri Kullanarak Web Uygulama Saldırılarının Tespitinde Genetik Öznitelik Seçimi Yaklaşımı. *Türkiye Bilişim Vakfı Bilgisayar Bilimleri ve Mühendisliği Dergisi*, *14*(2), 109–119.

[2] Thomas, V. E., Ugwu, C., & Onyejegbu, L. (2021). Comparative Analysis of Dimensionality Reduction Techniques on Datasets for Zero-Day Attack Vulnerability. *Journal of Software Engineering and Simulation*, *7*(8), 48–56.

[3] Yalçınkaya, M. A., & Küçüksille, E. (2021). Web Uygulama Sızma Testlerinde Kapsam Genişletme İşlemi İçin Metodoloji Geliştirilmesi Ve Uygulanması. *Süleyman Demirel Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, *25*(1), 16–27. https://doi.org/10.19113/sdufenbed.661867

[4] Abomhara, M., & Køien, G. M. (2015). Cyber Security and the Internet of Things: Vulnerabilities, Threats, Intruders and Attacks. *Journal of Cyber Security and Mobility*, *4*, 65–88. https://doi.org/10.13052/jcsm2245-1439.414

[5] Humayun, M., Niazi, M., Jhanjhi, N. Z., Alshayeb, M., & Mahmood, S. (2020). Cyber Security Threats and Vulnerabilities: A Systematic Mapping Study. *Arabian Journal for Science and Engineering*, *45*(4), 3171–3189. https://doi.org/10.1007/s13369-019-04319-2

[6] Khan, F., Ncube, C., Ramasamy, L. K., Kadry, S., & Nam, Y. (2020). A Digital Dna Sequencing Engine for Ransomware Detection Using Machine Learning. *IEEE Access*, *8*, 119710–119719. https://doi.org/10.1109/ACCESS.2020.3003785

[7] Khan, N. F., Ikram, N., Murtaza, H., & Asadi, M. A. (2021). Social Media Users and Cybersecurity Awareness: Predicting Self-Disclosure Using a Hybrid Artificial Intelligence Approach. *Kybernetes*. https://doi.org/10.1108/K-05-2021-0377

[8] Erhan, D., & Anarım, E. (2020). İstatistiksel Yöntemler ile DDoS Saldırı Tespiti DDoS Detection Using Statistical Methods. *2020 28th Signal Processing and Communications Applications Conference (SIU)*, 1–4.

[9] İlker, K. (2019). Kaba Kuvvet Saldırı Tespiti Ve Teknik Analizi. *Sakarya University Journal of Computer and Information Sciences*, *2*(2), 61–69. https://doi.org/10.35377/saucis.02.02.561844

[10] Khonji, M., Iraqi, Y., & Jones, A. (2013). Phishing Detection: A Literature Survey. *IEEE Communications Surveys & Tutorials*, *15*(4), 2091–2121. https://doi.org/10.1109/SURV.2013.032213.00009

[11] Özdemir, E., & Türkoğlu, İ. (2022). Yazılım Güvenlik Açıklarının Evrişimsel Sinir Ağları. *Fırat Üniversitesi Mühendislik Bilimleri Dergisi*, *34*(2), 517–529.

[12] *Wikipedia*. (2022). Wikipedia. https://tr.wikipedia.org/wiki/SQL_Enjeksiyonu

[13] Saad, M., Khormali, A., & Mohaisen, A. (2018). End-to-End Analysis of in-Browser Cryptojacking. *ArXiv Preprint ArXiv:1809.02152*.

[14] Ahmad, A., Shafiuddin, W., Kama, M. N., & Saudi, M. M. (2019). A New Cryptojacking Malware Classifier Model Based on Dendritic Cell Algorithm. *Proceedings of the 3rd International Conference on Vision, Image and Signal Processing*, 1–5.

[15] Sarker, I. H., Kayes, A. S. M., Badsha, S., Alqahtani, H., Watters, P., & Ng, A. (2020). Cybersecurity Data Science: An Overview from Machine Learning Perspective. *Journal of Big Data*, *7*(1), 1–29. https://doi.org/10.1186/s40537-020-00318-5

[16] Kılınçer, İ. F. (2022). *Saldırı Tespit Ve Engelleme Sistemleri İçin Yapay Zeka Tabanlı Yeni Bir Güvenlik Modelinin Oluşturulması* [Doktora Tezi]. Fırat Üniversitesi, Fen Bilimleri Enstitüsü.

[17] Hindy, H., Atkinson, R., Tachtatzis, C., Colin, J.-N., Bayne, E., & Bellekens, X. (2020). Utilising Deep Learning Techniques for Effective Zero-Day Attack Detection. *Electronics*, *9*(10), 1684. https://doi.org/10.3390/electronics9101684

[18] Metlek, S. (2021). Disease Detection from Cassava Leaf Images with Deep Learning Methods in Web Environment. *International Journal of 3D Printing Technologies and Digital Industry*, *5*(3), 625–644.

[19] Metlek, S. (2022). Forecasting of Dow Jones Sukuk Index Prices Using Artificial Intelligence Systems. *Economic Computation & Economic Cybernetics Studies & Research*, *56*(1).

[20] Hindy, H., Atkinson, R., Tachtatzis, C., Colin, J.-N., Bayne, E., & Bellekens, X. (2020). Utilising Deep Learning Techniques for Effective Zero-Day Attack Detection. *Electronics*, *9*(10), 1684.

[21] Al-rimy, B. A. S., Maarof, M. A., Prasetyo, Y. A., Shaid, S. Z. M., & Ariffin, A. F. M. (2018). Zero-Day Aware Decision Fusion-Based Model for Crypto-Ransomware Early Detection. *International Journal of Integrated Engineering*, *10*(6). https://publisher.uthm.edu.my/ojs/index.php/ijie/article/view/2828

[22] Sharma, V., Kim, J., Kwon, S., You, I., Lee, K., & Yim, K. (2018). A Framework for Mitigating Zero-Day Attacks in Iot. *ArXiv Preprint ArXiv:1804.05549*.

[23] Sun, X., Dai, J., Liu, P., Singhal, A., & Yen, J. (2018). Using Bayesian Networks for Probabilistic Identification of Zero-Day Attack Paths. *IEEE Transactions on Information Forensics and Security*, *13*(10), 2506–2521.

[24] Zhou, Q., & Pezaros, D. (2019). Evaluation of Machine Learning Classifiers for Zero-Day Intrusion Detection–An Analysis on CIC-AWS-2018 dataset. *ArXiv Preprint ArXiv:1905.03685*.

[25] Zhao, J., Shetty, S., Pan, J. W., Kamhoua, C., & Kwiat, K. (2019). Transfer Learning for Detecting Unknown Network Attacks. *EURASIP Journal on Information Security*, *2019*(1), 1–13.

[26] Obeidat, I., Hamadneh, N., Alkasassbeh, M., Almseidin, M., & AlZubi, M. (2019). *Intensive Pre-Processing of Kdd Cup 99 for Network Intrusion Classification Using Machine Learning Techniques*.

[27] Blaise, A., Bouet, M., Conan, V., & Secci, S. (2020). Detection of Zero-Day Attacks: An Unsupervised Port-Based Approach. *Computer Networks*, *180*, 107391. https://doi.org/10.1016/j.comnet.2020.107391

[28] Sameera, N., & Shashi, M. (2020). Deep Transductive Transfer Learning Framework for Zero-Day Attack Detection. *ICT Express*, *6*(4), 361–367.

[29] Nkongolo, M., Van Deventer, J. P., & Kasongo, S. M. (2021). UGRansome1819: A Novel Dataset for Anomaly Detection and Zero-Day Threats. *Information*, *12*(10), 405.

[30] Nkongolo, M., Van Deventer, J. P., Kasongo, S. M., Zahra, S. R., & Kipongo, J. (2022). A Cloud Based Optimization Method for Zero-Day Threats Detection Using Genetic Algorithm and Ensemble Learning. *Electronics*, *11*(11), 1749.

[31] Nkongolo, M., van Deventer, J. P., & Kasongo, S. M. (2022). The Application of Cyclostationary Malware Detection Using Boruta and PCA. In *Computer Networks and Inventive Communication Technologies* (pp. 547–562). Springer.

[32] Aygün, R. C. (2017). *Derin Öğrenme Yöntemleri İle Bilgisayar Ağlarında Güvenliğe Yönelik Anormallik Tespiti* [PhD Thesis]. Yıldız Teknik Üniversitesi, Fen Bilimleri Enstitüsü.

[33] Tavallaee, M., Bagheri, E., Lu, W., & Ghorbani, A. A. (2009). A Detailed Analysis of the Kdd Cup 99 Data Set. *2009 IEEE Symposium on Computational Intelligence for Security and Defense Applications*, 1–6.

[34] Zhou, Z.-H., & Feng, J. (2019). Deep Forest. *National Science Review*, *6*(1), 74–86.

[35] Guo, Y., Liu, S., Li, Z., & Shang, X. (2018). Bcdforest: A Boosting Cascade Deep Forest Model Towards the Classification of Cancer Subtypes Based on Gene Expression Data. *BMC Bioinformatics*, *19*(5), 1–13.

[36] Weinstein, B. G., Marconi, S., Bohlman, S. A., Zare, A., & White, E. P. (2019). Geographic Generalization in Airborne Rgb Deep Learning Tree Detection. *BioRxiv*, 790071.

COMBINING GREY WOLF OPTIMIZATION WITH HARRIS HAWK OPTIMIZATION FOR BINARY FEATURE SELECTION

Onur INAN [1] and Mustafa Serter UZER [2]

INTRODUCTION

In order to reduce the processing load of the computer systems used and to increase the success of computer software, important, relevant and necessary data must be selected first and then processed. Considering that so much information is generated from a wide variety of devices, how necessary this is unavoidable. Moreover, thanks to the devices and methods developed day by day, the production of data in many different fields is increasing. The data preprocessing step, also known as feature selection or feature subset selection, means selecting features that are more meaningful from the data, that is, that contribute more to the classification [1]. To give an example of medical data, the selection of relevant features for the diagnosis of various diseases can both increase the success of disease diagnosis and shorten the diagnosis time. Reducing data in other fields can contribute by increasing the classification success in that field and shortening the classification processing time.

A feature selection algorithm consists of a search technique and evaluation measure. New feature subsets are proposed using search algorithms, and the various feature subsets are scored using evaluation metrics [1, 2]. The most straightforward technique is testing each potential subset of attributes to choose the one that minimizes the error rate. With the exception of the lowest feature sets, this is an exhaustive search of the space and is computationally intractable. The algorithm is greatly influenced by the evaluation metric that is used, and it is these evaluation metrics that distinguish between the three main types of feature selection algorithms: wrappers, filters, and embedding techniques. Due to the fact that it necessitates learning and assessing the classifier in order to choose a subset of candidates, the wrapper technique typically produces better results than filtering methods [3]. Wrapping techniques often produce the most ideal feature set, but they are computationally demanding compared to other approaches since they require new training for each subset. There is a measure in filter methods. This measure uses the feature subset to score. The feature set with the highest utility is often chosen in

1   Onur Inan, Dr, Department of Computer Engineering, Faculty of Technology, Selcuk University, Turkey, ORCID No: https://orcid.org/0000-0003-4573-7025
2   Mustafa Serter Uzer, Dr, Department of Electronics and Automation, Ilgın Vocational School, Selcuk University, Turkey, ORCID No: https://orcid.org/0000-0002-8829-5987

accordance with this measure, which is typically chosen in a way that will speed up the subset evaluation. In general, filter approaches need less computing power and perform less accurately in terms of prediction than a wrapper. Embedded methods refers to a broad class of methods that execute feature selection throughout the model development process and in terms of computing complexity, these methods often fall between filters and wrappers. Embedded techniques do variable selection in the training process and embedded techniques are frequently unique to a particular learning machine [4].

Optimization techniques called Harris Hawk Optimization (HHO) and Grey Wolf Optimization (GWO) draw inspiration from nature [5, 6]. It is made use of meta-heuristic methods such as Simulated Annealing (SA), Particle Swarm Optimization (PSO), Tabu Search, Ant Colony Optimization (ACO), Genetic Algorithm (GA), and Artificial Neural Networks (ANN) in the literature. Furthermore, Bee Colony Algorithm (BCA), Harmonic Search Algorithm, Ant Lion Optimizer (ALO), and Kangaroo Algorithm are more contemporary or less investigated meta-heuristics methods [7-9]. Optimization techniques are used for feature selection. Some of these are methods in which optimization is used alone, while others are methods that are performed by hybridizing optimization techniques. Some of these are presented in this section. A hybrid technique for feature selection called WOASAT-2 is proposed in [9] that combines the WOA, Simulated Annealing (SA) algorithm, and Tournament Selection method. The binary-based hybrid GWO and PSO technique (BGWOPSO) [10] and the binary GWO method (bGWO) made up of two approaches [11] were presented for feature selection.

In this work, a hybrid binary optimization approach made up of GWO and HHO has been used to build a more effective feature selection method. For the purpose of evaluating the suggested approach, five data sets [12] were employed. Sections 2-4 include explanations of the GWO, HHO, and LDA procedures, respectively. Section 5 explains the suggested approach. In Section 6, findings from the created FSGWOHHO technique are presented, along with a comparison to findings from previous literature investigations. Section 7 provides the suggested method's conclusions.

GREY WOLF OPTIMIZATION (GWO)

Mirjalili et al. devised the GWO optimization approach based on the hunting strategies that grey wolves developed in accordance with their social hierarchy [6]. As seen in Figure 1, grey wolves can be classified into four groups based on their hierarchical positions: alpha(α), beta(β), delta(δ), and omega(ω). Important decisions like hunting are made by the alpha group, which consists of a female and a male wolf. The "beta" or second-ranked wolves in the hierarchy aid the alpha wolf in making decisions. Delta wolves are the third-ranked wolves in the pack and adhere to the alpha and beta wolves' orders. Omegas are the last to be allowed to consume because they are at the bottom of the wolf food chain.
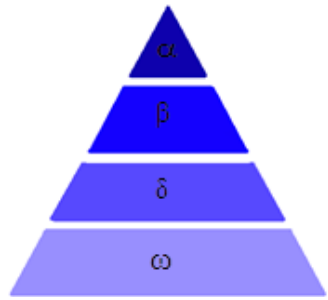
Figure 1. Diagram showing the social hierarchy of grey wolves

Encircling the prey, hunting, attacking the prey (exploitation), and seeking out prey are the four stages of the GWO method. Grey wolves encircling prey is represented in Equations (1) and (2). Grey wolves are located at $\vec{H}(t+1)$, while the prey's location at iteration i is represented by $\vec{H}l(t)$.

$$\vec{H}(t+1)=\vec{H}l(t)-\vec{A}.\vec{D} \tag{1}$$

$$\vec{D}=\left|\vec{C}.\vec{H}l(t)-\vec{H}(t)\right| \tag{2}$$

$$\vec{A}=2\vec{a}.\vec{r_1}+\vec{a} \tag{3}$$

$$\vec{C}=2.\vec{r_2} \tag{4}$$

The coefficient vectors A and C are produced from Equations (3) and (4). The [0, 1] range of random values make up the $\vec{r_1}$ and $\vec{r_2}$ variables. The a value drops linearly from 2 to 0. By changing the A and C vectors, the closest new locations of the best agent may be tested. Following the surrounding phase, the hunting phase begins the search for the optimal response. During this phase, alpha controls hunting, with occasional participation from beta and delta. The three best places are therefore used to update the location of the grey wolves using Equations (5), (6), and (7).

$$\vec{D_\alpha}=\left|\vec{C_1}.\vec{H_\alpha}-\vec{H}\right|, \quad \vec{D_\beta}=\left|\vec{C_2}.\vec{H_\beta}-\vec{H}\right|, \quad \vec{D_\delta}=\left|\vec{C_3}.\vec{H_\delta}-\vec{H}\right| \tag{5}$$

$$\vec{H_1}=\vec{H_\alpha}-\vec{A_1}.\vec{D_\alpha}, \quad \vec{H_2}=\vec{H_\beta}-\vec{A_2}.\vec{D_\beta}, \quad \vec{H_3}=\vec{H_\delta}-\vec{A_3}.\vec{D_\delta} \tag{6}$$

$$\vec{H}(t+1)=\frac{\vec{H_1}+\vec{H_2}+\vec{H_3}}{3} \tag{7}$$

The GWO algorithm, in summary, creates a haphazard population. Alpha, beta, and delta wolves can all estimate the location of probable prey. Next, the distance of the potential solution is revised. Then, from 2 to 0, a is lowered to emphasize exploration and exploitation, respectively. They move to the prey if $A<1$. If $A>1$, they

cease their attack on the prey. GWO comes to an end once it has been satisfactorily resolved [6].The pseudocode for the GWO algorithm is shown in Figure 2.

```
Initialize the grey wolf population Hᵢ (i = 1, 2, ..., n)
Initialize a, A, C
Calculate the fitness of each search agent
H_α= The best search agent
H_β= The second best search agent
H_δ= The third best search agent
   while (t < Maximum iteration number)
       for each search agent
           Update the position of the current search agent using Eq (7)
       end for
   Update a, A, C
   Calculate the fitness of all search agent
   Update H_α, H_β, H_δ
   t=t+1
   end while
return H_α
```

Figure 2. Pseudocode of the GWO algorithm [6]

HARRIS HAWKS OPTIMIZATION (HHO)

The HHO algorithm is an optimization method that imitates the coordinated action of numerous hawks striking at prey from different angles to confuse it [5]. In response to the escape mechanisms used by their victims, hawks have evolved a range of hunting strategies. The mathematical modeling of these various strategies leads to the creation of HHO. The exploring a prey, surprise pounce, and varied assault strategies used by Harris hawks in are used to simulate the exploration and exploitation phases of HHO[5].

Exploration Phase

In HHO, the Harris' hawks perch in various locations and wait for prey to show up utilizing one of two strategies in Equation (8).

$$H(t+1)=\begin{cases} H_{rand}(t)-r_1\left|H_{rand}(t)-2r_2H(t)\right| & q \geq 0.5 \\ (H_{rabbit}(t)-H_m(t))-r_3(LB+r_4(UB-LB)) & q < 0.5 \end{cases} \quad (8)$$

While $H_{rand}(t$ is a hawk chosen at random from the current population, H_m is the mean location of the current hawk population. The vector of the rabbit's position is $H_{rabbit}(t)$, while the vector of the hawks' current position is $H(t)$. The variables *r1, r2, r3, r4*, and are updated and contain random numbers in each iteration (0,1). In *LB* and *UB*, the variables' upper and lower bounds are shown. One can determine the hawks' normal position using Equation (9).

$$H_m(t) = \frac{1}{N}\sum_{i=1}^{N} H_i(t) \tag{9}$$

where the total number of hawks is N. Additionally, $H_i(t)$ represents each hawk's location in iteration t.

Exploration to Exploitation Transition

The energy of the rabbit is described as follows using Equation 10.

$$E = 2E_0\left(1 - \frac{t}{T}\right) \tag{10}$$

where T is the number of permitted iterations. Its initial energy state is E_0, while the prey's escape energy is E.

Exploitation Phase

There are four different types of besiege.

Soft Besiege

The following Equations (11–12) can be used to define soft besiege.

$$H(t+1) = \Delta H(t) - E\left|JH_{rabbit}(t) - H(t)\right| \tag{11}$$

$$\Delta H(t) = H_{rabbit}(t) - H(t) \tag{12}$$

the randomized leap power of the rabbit when it is running away is $J=2(1-r_5)$. Moreover, r_5 denotes a random number made up: (0,1). $H(t)$ denotes the difference between the rabbit's current location and its position vector during cycle t. The J value varies at random throughout each iteration to reflect the characteristics of rabbit movement.

Hard Besiege

Equation (13) is used to update the current positions in a hard besiege scenario:

$$H(t+1) = H_{rabbit}(t) - E\left|\Delta H(t)\right| \tag{13}$$

Soft Besiege with Progressive Rapid Dives

It is hypothesized that the hawks could decide to launch a soft besiege based on the following rule in Equation (14) to choose their next course of action.

$$Y = H_{rabbit}(t) - E\left|JH_{rabbit}(t) - H(t)\right| \tag{14}$$

They are anticipated to dive in accordance with the rule described by Equation (15) based on the predictions made using LF.

$$Z = Y + S \times LF(D) \tag{15}$$

where LF is the levy flight function, which is obtained from Equation (16), D is the dimension of the problem, and S is a random vector with size $1xD$.

$$LF(h) = 0.01 \times \frac{u \times \sigma}{|v|^{\frac{1}{\beta}}}, \quad \sigma = \left\{ \frac{\Gamma(1+\beta) \times \sin(\frac{\pi\beta}{2})}{\Gamma(\frac{1+\beta}{2}) \times \beta \times 2^{(\frac{\beta-1}{2})}} \right\}^{\frac{1}{\beta}} \tag{16}$$

where u and v are arbitrary values between (0, 1) and 1.5 is used as the default constant β. Therefore, during the soft besiege phase, Equation (17) can be utilized to update the hawks' positions.

$$H(t+1) = \begin{cases} Y & if \ F(Y) < F(H(t)) \\ Z & if \ F(Z) < F(H(t)) \end{cases} \tag{17}$$

Here, Equations (14-15) are used to determine Y and Z.

Hard Besiege with Progressive Rapid Dives

Here, Equation (18)'s following rule is applicable:

$$H(t+1) = \begin{cases} Y & if \ F(Y) < F(H(t)) \\ Z & if \ F(Z) < F(H(t)) \end{cases} \tag{18}$$

where the new rules in Equations (19) and (20) are used to derive Y and Z. The pseudocode for HHO is shown in Figure 3.

$$Y = H_{rabbit}(t) - E|JH_{rabbit}(t) - H_m(t)| \tag{19}$$
$$Z = Y + S \times LF(D) \tag{20}$$

Inputs: The population size N and maximum number of iterations T
Outputs: The location of rabbit and its fitness value

Initialize the random population $H_i(i = 1, 2, \ldots, N)$
while (stopping condition is not met) **do**
　　　　Calculate the fitness values of hawks
　　Set H_{rabbit} as the location of rabbit (best location)
　　for (each hawk (H_i)) **do**
　　　　Update the initial energy E_0 and jump strength J
　　　　Update the E using Eq. (10)

　　　　if (|E|≥ 1) **then**
　　　　　　Update the location vector using Eq. (8)

　　　　if (|E|< 1) **then**

　　　　　　if (r ≥0.5 and |E|≥ 0.5) **then**
　　　　　　　　Update the location vector using Eq. (11)

　　　　　　else if (r ≥0.5 and |E|< 0.5) **then**
　　　　　　　　Update the location vector using Eq. (13)

　　　　　　else if (r < 0.5 and |E|≥ 0.5) **then**
　　　　　　　　Update the location vector using Eq. (17)

　　　　　　else if (r < 0.5 and |E|< 0.5) **then**
　　　　　　　　Update the location vector using Eq. (18)

Return H_{rabbit}

Figure 3. Pseudocode of the HHO [5]

LINEAR DISCRIMINANT ANALYSIS (LDA)

Data categorization and dimensionality reduction are done using LDA. The LDA seeks a linear combination of characteristics that can distinguish across classes [13]. LDA is utilized twice in this study for two distinct objectives. It is first applied to the feature selection optimization algorithm's objective function. Second, it was applied to evaluate the accuracy of the chosen characteristics' categorization.

$$S_b = \frac{1}{N} \sum_{i=1}^{c} N_i (\overline{x_i} - \overline{x})(\overline{x_i} - \overline{x})^T$$

$$S_W = \frac{1}{N} \sum_{i=1}^{c} \sum_{s=1}^{N_i} (x_s^i - \overline{x_i})(x_s^i - \overline{x_i})^T \tag{21}$$

In Equation (21), the between-class scatter matrix S_b and the within-class scatter matrix S_w are both obtained [14]. LDA attempts to address the optimization problem, which is provided in Equation (22) [14].

$$\max_{w} \frac{tr(W^T S_b W)}{tr(W^T S_w W)} \tag{22}$$

where W is optimal projection transformation matrix. W is maximized the ratio of between-class scatter to within-class scatter in the projected space of $W \in R^{n \times d}$ (d $\leq$ n) [14].

THE PROPOSED METHOD (FSGWOHHO)

To choose more pertinent features that would improve the classification's accuracy, a binary hybrid technique using GWO and HHO is suggested. This newly created feature selection technique is named as FSGWOHHO. While creating a new population element in the pseudocodes of both GWO and HHO, conversion to binary format is achieved using Equation (23). Thus, the number 1 means the selected feature and the number 0 means the feature is not selected.

In order for the position update to work in binary space, it can be transformed into Equation (23) [10, 11].

$$x_d^{t+1} = \begin{cases} 1 & if\ sigmoid(x_1) \geq rand \\ 0 & otherwise \end{cases} \tag{23}$$

Iteration t's location in dimension d, x_d^{t+1}, is changed to either 0 or 1 in Equation (23). Randomly selected between 0 and 1, $rand$ is a number. The $sigmoid(a)$ is provided in Equation (24) [11]:

$$sigmoid(a) = \frac{1}{1 + e^{-10(x-0.5)}} \tag{24}$$

The suggested FSGWOHHO's binary optimization algorithm's fitness function is provided in Equation (25).

$$fitness = \alpha\rho_R(D) + \beta\frac{|S|}{|T|} \tag{25}$$

where $\rho_R(D)$ is the LDA classifier's error rate, $|T|$ is the total number of features in the dataset, and $|S|$ is the total number of features that were chosen. β is a parameter that is equal to $(1-\alpha)$, while α is a parameter having a value between 0 and 1[10, 11]. In our application, β is assumed to be 0.01. In the suggested strategy, there are 10 search agents, and 100 iterations are the maximum allowed. Figure 4 contains the proposed FSGWOHHO's pseudocode.

```
Create random binary population Xₖ(k = 1, 2, . . . , N)
r=0
while (r<T) do
    [Alpha_score, Alpha_pos, gPositions]=bGWO(Positions, fobj, r);
    [Rabbit_Energy, Rabbit_Energy_pos, hPositions]=bHHO(Positions, fobj, r);
    if Alpha_score < Rabbit_Energy
    Best_score = Alpha_score;
    Best_pos = Alpha_pos;
    Positions=gPositions;
    else
    Best_score = Rabbit_Energy;
    Best_pos = Rabbit_Energy_pos;
    Positions=hPositions;
    end if
    r=r+1
end while
Return Best  score, Best  pos
```

Figure 4. Pseudocode of the proposed FSGWOHHO

The FSGWOHHO technique used five datasets to identify the key features. For a more precise categorization, the LDA classifier was given a dataset with these characteristics. The ratio of training to testing was chosen at 50% to allow comparison with other research using swarm optimization techniques in the literature.

THE PROPOSED METHOD RESULTS AND DISCUSSION

Table 1 displays the attributes of the BreastEW, WineEW, Breastcancer, HeartEW and Lymphography datasets [12] utilized in the FSGWOHHO technique. The average outcomes from ten runs of the proposed FSGWOHHO technique and categorization were compared to the existing literature. Table 2 compares the average classification accuracies (ACA). In terms of ACA, it has been observed that the suggested FSGWOHHO achieves better than the literature research.

Table 1. Properties of the data sets used for the FSGWOHHO method

Dataset	# Features	# Instances
BreastEW	30	569
WineEW	13	178
Breastcancer	9	699
HeartEW	13	270
Lymphography	18	148

Table 2. The comparison of the ACA between the proposed FSGWOHHO and literature studies

Dataset	PSO [9]	ALO [9]	GA [9]	WOASAT-2 [9]	FSGWOHHO
BreastEW	0.94	0.93	0.94	0.98	0.99
WineEW	0.95	0.91	0.93	0.99	1.00
Breastcancer	0.95	0.96	0.96	0.97	0.98
HeartEW	0.78	0.83	0.82	0.85	0.89
Lymphography	0.69	0.79	0.71	0.89	0.91

The comparison of the literature in terms of ACA for BreastEW dataset, for WineEW dataset, for Breastcancer dataset, for HeartEW dataset, for Lymphography dataset are given in Figure 5, Figure 6, Figure 7, Figure 8 and Figure 9, respectively.
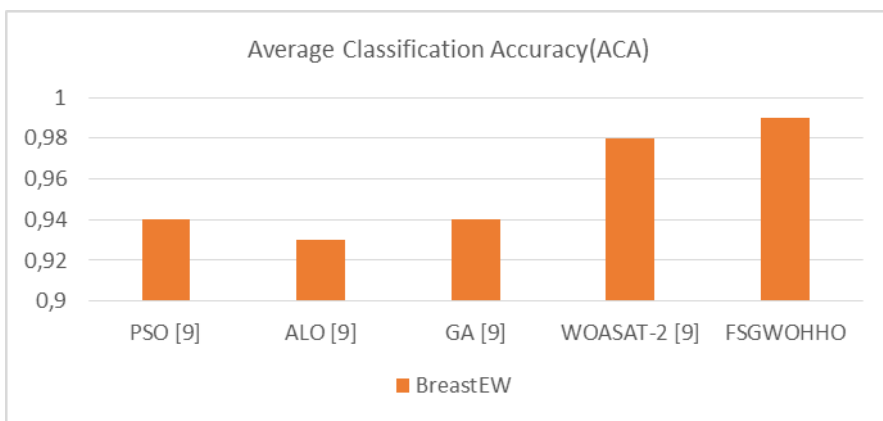
Figure 5. The comparison of the literature for BreastEW dataset

For BreastEW data, the best classification accuracy (BCA) is the proposed FSGWOHHO method, while the closest classification accuracy is the WOASAT-2 method. Also, the worst classification accuracy (WCA) is ALO.
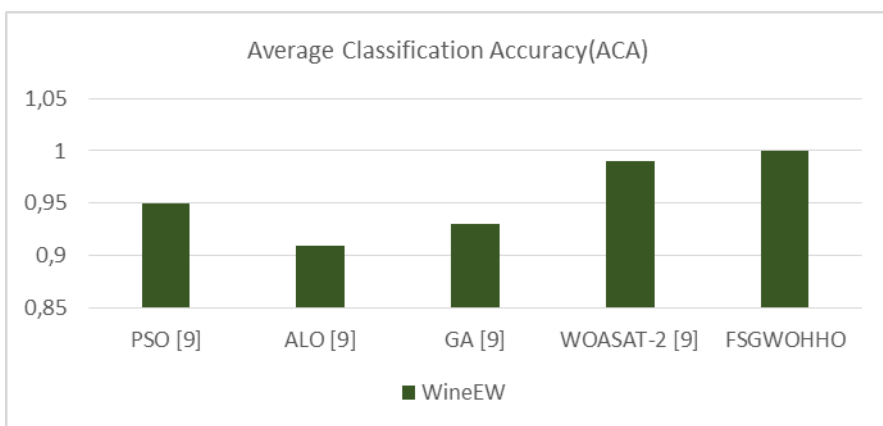
Figure 6. The comparison of the literature for WineEW dataset

For WineEW data, the BCA is the proposed FSGWOHHO method, while the closest classification accuracy is the WOASAT-2 method. Also, the WCA is ALO.
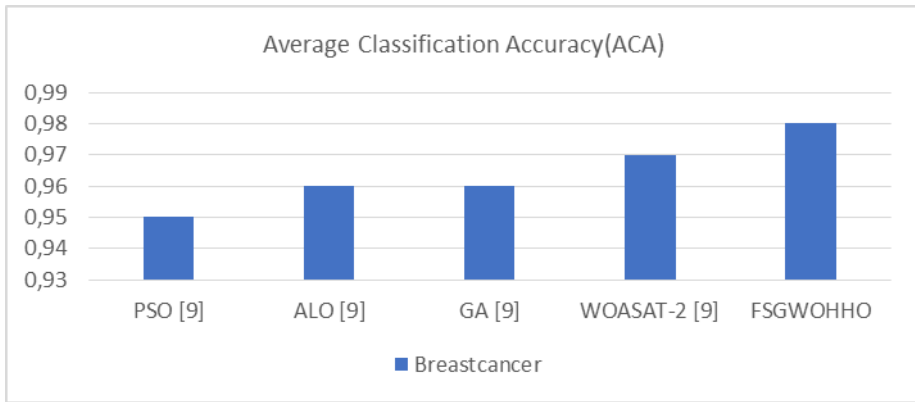
Figure 7. The comparison of the literature for Breastcancer dataset

For Breastcancer data, the BCA is the proposed FSGWOHHO method, while the closest classification accuracy is the WOASAT-2 method. Also, the WCA is PSO.

Figure 8. The comparison of the literature for HeartEW dataset

For HeartEW data, the BCA is the proposed FSGWOHHO method, while the closest classification accuracy is the WOASAT-2 method. Also, the WCA is PSO.
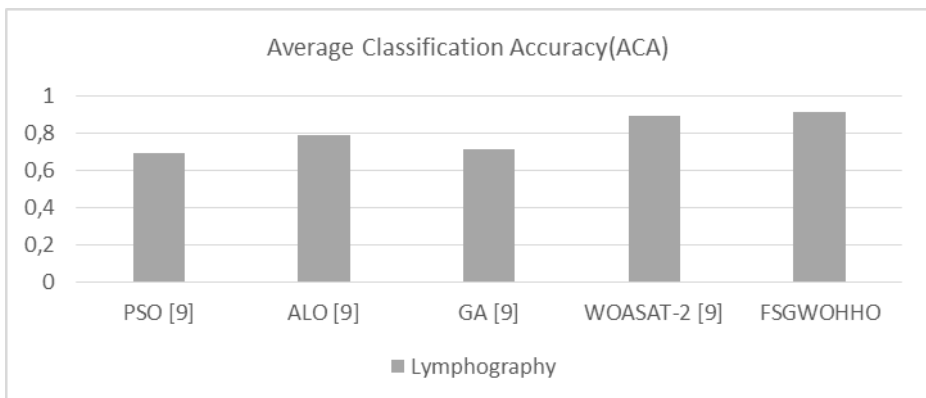
Figure 9. The comparison of the literature for Lymphography dataset

For Lymphography data, the BCA is the proposed FSGWOHHO method, while the closest classification accuracy is the WOASAT-2 method. Also, the WCA is GA.

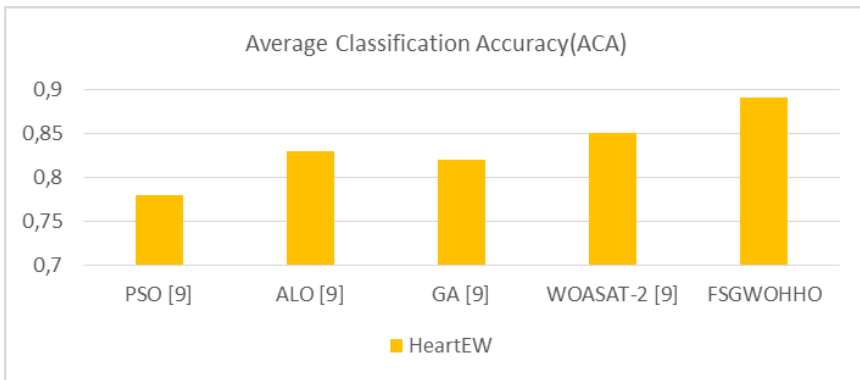Table 3-5 compares the worst fitness functions (WFF), average fitness functions AFF), and best fitness functions (BFF). The literature comparisons in terms of the WFF, AFF, and BFF are given in Figure 10-12. Table 6 compares the average counts of the attributes that were chosen.

Table 3. The WFF comparisons between the proposed FSGWOHHO and literature studies

Dataset	PSO [9]	ALO [9]	GA [9]	WOASAT-2 [9]	FSGWOHHO
BreastEW	0.05	0.04	0.05	0.04	0.02
WineEW	0.03	0.03	0.03	0.03	0.05
Breastcancer	0.03	0.03	0.04	0.04	0.03
HeartEW	0.18	0.13	0.14	0.18	0.12
Lymphography	0.27	0.16	0.27	0.14	0.15
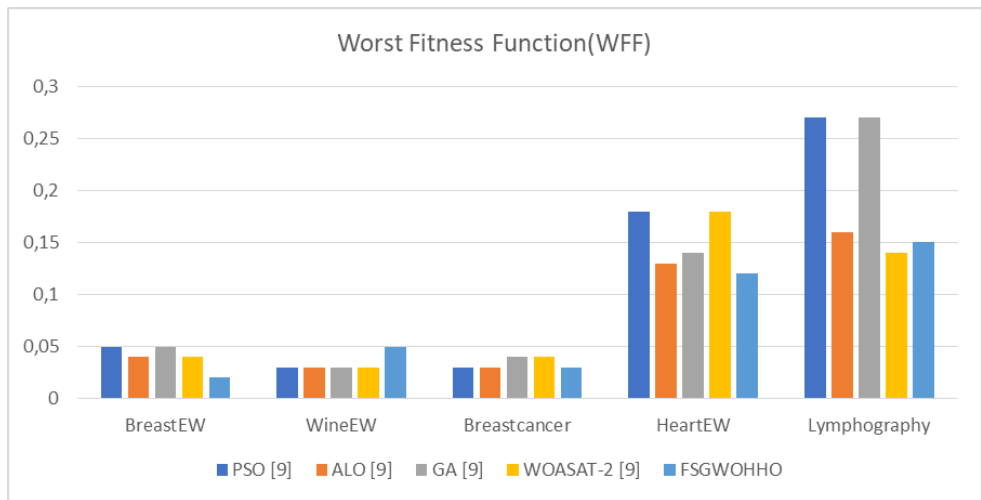
Figure 10. The literature comparison in terms of the WFF

Except for the WineEW dataset, the WFF value is the lowest in the proposed FSGWOHHO method, while for the WineEW dataset, the WFF value is the lowest in all other algorithms.

Table 4. AFF comparisons between the proposed FSGWOHHO and literature studies

Dataset	PSO [9]	ALO [9]	GA [9]	WOASAT-2 [9]	FSGWOHHO
BreastEW	0.03	0.03	0.04	0.03	0.02
WineEW	0.02	0.01	0.02	0.01	0.01
Breastcancer	0.03	0.02	0.03	0.04	0.03
HeartEW	0.15	0.12	0.14	0.16	0.11
Lymphography	0.19	0.14	0.17	0.11	0.11
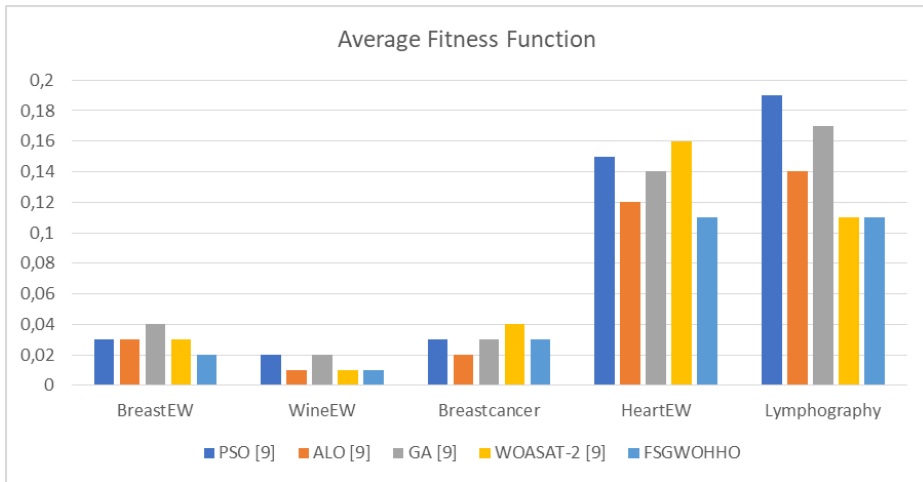
Figure 11. The literature comparison in terms of the AFF

Except for the Breastcancer dataset, the AFF value is either the same or higher in the proposed FSGWOHHO method, while for the Breastcancer dataset, the AFF value is the lowest in ALO.

Table 5. The BFF comparisons between the proposed FSGWOHHO and literature studies

Dataset	PSO [9]	ALO [9]	GA [9]	WOASAT-2 [9]	FSGWOHHO
BreastEW	0.02	0.03	0.02	0.02	0.01
WineEW	0.00	0.00	0.00	0.00	0.00
Breastcancer	0.03	0.02	0.02	0.03	0.02
HeartEW	0.13	0.11	0.12	0.13	0.09
Lymphography	0.14	0.08	0.12	0.09	0.06

FSGWOHHO method provided either the same or higher BFF value for all datasets compared to other methods. In addition, the FSGWOHHO method provided the lowest average selected feature number for BreastEW and WineEW datasets. WOASAT-2 method provided the lowest average selected feature number for other datasets.
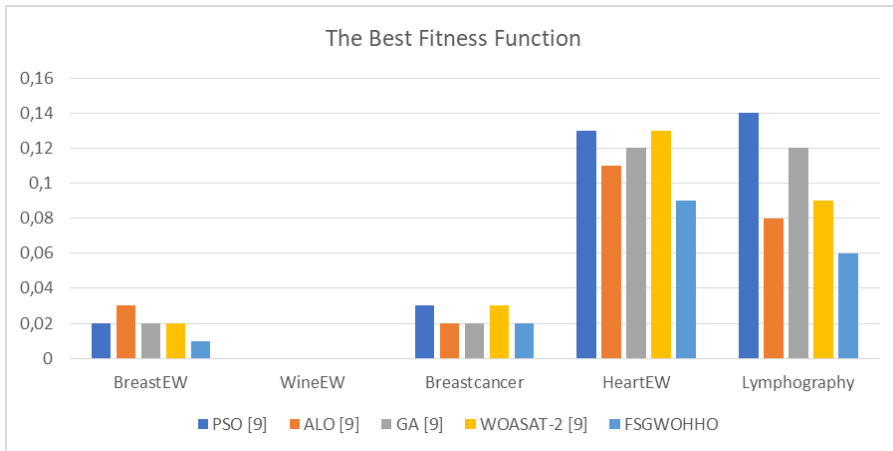
Figure 12. The literature comparison in terms of the BFF

Table 6. The comparison of average numbers of selected feature between the proposed FSGWOHHO and literature studies

Dataset	PSO [9]	ALO [9]	GA [9]	WOASAT-2 [9]	FSGWOHHO
BreastEW	**16.56**	**16.08**	**16.35**	**11.60**	**9.9**
WineEW	**8.36**	**10.70**	**8.63**	**6.40**	**5.6**
Breastcancer	**5.72**	**6.28**	**5.09**	**4.2**	**4.8**
HeartEW	**7.94**	**10.31**	**9.49**	**5.4**	**7.0**
Lymphography	**8.98**	**11.05**	**11.05**	**7.2**	**7.8**

CONCLUSION

In this work, the widely used optimization techniques GWO and HHO were combined to generate the more effective binary hybrid feature selection approach FSGWOHHO. Five UCI datasets have been used to test the proposed technique. By using LDA, selected features are classified. Ten repetitions were run through the suggested feature selection categorization algorithm. The average number of features, average classification successes, worst, average, and best fitness function values were determined for the suggested feature selection. The average categorization success rate for BreastEW was 99%, for WineEW was 100%, for Breastcancer was 98%, for HeartEW was 89%, and for Lymphography was 91%. For each dataset, pertinent comparison graphs are created. The number of features in the data collection is cut in half with the suggested feature selection because extraneous or pointless features are eliminated. The proposed method's features were more successful at classifying data. The results were found to be encouraging when they were compared to the literature.

ACKNOWLEDGMENTS

REFERENCES

[1] M. S. Uzer, O. Inan, and N. Yılmaz, "A hybrid breast cancer detection system via neural network and feature selection based on SBS, SFS and PCA," *Neural Computing and Applications,* vol. 23, no. 3, pp. 719-728, 2013/09/01 2013, doi: 10.1007/s00521-012-0982-6.

[2] M. S. Uzer, N. Yilmaz, and O. Inan, "Feature Selection Method Based on Artificial Bee Colony Algorithm and Support Vector Machines for Medical Datasets Classification," *The Scientific World Journal,* vol. 2013, p. 419187, 2013/07/28 2013, doi: 10.1155/2013/419187.

[3] R. Kohavi and G. H. John, "Wrappers for feature subset selection," (in English), *Artif Intell,* vol. 97, no. 1-2, pp. 273-324, Dec 1997. [Online]. Available: <Go to ISI>://WOS:000071321500009.

[4] I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," *Journal of machine learning research,* vol. 3, no. Mar, pp. 1157-1182, 2003.

[5] A. A. Heidari, S. Mirjalili, H. Faris, I. Aljarah, M. Mafarja, and H. L. Chen, "Harris hawks optimization: Algorithm and applications," *Future Generation Computer Systems-the International Journal of Escience,* vol. 97, pp. 849-872, Aug 2019, doi: 10.1016/j.future.2019.02.028.

[6] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey Wolf Optimizer," (in English), *Advances in Engineering Software,* vol. 69, pp. 46-61, Mar 2014. [Online]. Available: <Go to ISI>://WOS:000331924800006.

[7] A. Karadeniz and Y. Çelik, "Whale Optimization Algorithm for Numerical Constrained Optimization," *Academic Platform Journal of Engineering and Science,* vol. 8, no. 3, pp. 547-554.

[8] M. C. Akkoyunlu and O. Engin, "Kesikli Harmoni Arama Algoritması İle Optimizasyon Problemlerinin Çözümü: Literatür Araştırması," *Selçuk Üniversitesi Mühendislik, Bilim ve Teknoloji Dergisi,* vol. 26, no. 4, pp. 140-148, 2011.

[9] M. M. Mafarja and S. Mirjalili, "Hybrid Whale Optimization Algorithm with simulated annealing for feature selection," (in English), *Neurocomputing,* vol. 260, pp. 302-312, Oct 18 2017. [Online]. Available: <Go to ISI>://WOS:000405536900032.

[10] Q. Al-Tashi, S. J. A. Kadir, H. M. Rais, S. Mirjalili, and H. Alhussian, "Binary Optimization Using Hybrid Grey Wolf Optimization for Feature Selection," (in English), *Ieee Access,* vol. 7, pp. 39496-39508, 2019. [Online]. Available: <Go to ISI>://WOS:000463930300001.

[11] E. Emary, H. M. Zawba, and A. E. Hassanien, "Binary grey wolf optimization approaches for feature selection," (in English), *Neurocomputing,* vol. 172, pp. 371-381, Jan 8 2016. [Online]. Available: <Go to ISI>://WOS:000364884700036.

[12] C. L. Blake and C. J. Merz. "UCI repository of machine learning databases, 1998." http://www.ics.uci.edu/~mlearn/MLRepository.html (accessed 2012).

[13] K. Al-Dulaimi, V. Chandran, K. Nguyen, J. Banks, and I. Tomeo-Reyes, "Benchmarking HEp-2 specimen cells classification using linear discriminant analysis on higher order spectra features of cell shape," (in English), *Pattern Recogn Lett,* vol. 125, pp. 534-541, Jul 1 2019. [Online]. Available: <Go to ISI>://WOS:000482374500073.

[14] C. N. Li, Y. H. Shao, Z. Wang, N. Y. Deng, and Z. M. Yang, "Robust Bhattacharyya bound linear discriminant analysis through an adaptive algorithm," (in English), *Knowl-Based Syst,* vol. 183, Nov 1 2019. [Online]. Available: <Go to ISI>://WOS:000494891600005.

PRINT ATTACK DETECTION FOR EAR BIOMETRICS WITH FUSION OF TEXTURE-BASED AND CNN-BASED APPROACHES

İmren YEŞILYURT[1] , Mehtap Köse ULUKÖK[2] , Önsen TOYGAR[3]

INTRODUCTION

Trying to fool automated person identification systems by applying illegal ways is called spoofing. In the biometric community, some of the types of attacks can be applied to fool the identification systems [1] [2]. Detecting these types of attacks is called spoof detection in the biometric community. Researchers have been trying to find a robust method to defeat counterfeit attacks against biometric recognition systems that are based on face, iris, finger vein, palm vein, fingerprint, palmprint, audio and ear biometrics [3] [4] [5] [6] [7] [8] [9] [10] [11] [12] [13]. Beside, since ear biometric trait is unique and remains unchangeable during age variation, there are many identification systems based on ear biometric trait in the literature [14] [15] [16] [17]. However, there is no such a study that employs texture-based and deep learning methods which can detect attacks to ear based identification systems in the literature.

The method that is the fusion of deep learning approach and texture-based approach in this paper intends to overcome the attacks to ear recognition systems. Deep learning is a form of supervised learning. There are multiple levels where different feature representations are obtained in each level in deep learning methods. The most prominent advantage of deep learning is that a feature representation is learned from data naturally by employing a learning method, instead of handcrafted learning by engineers. Deep learning is applied successfully in many applications such as medical image analysis, human action recognition, autonomous driving, text classification, face recognition, spoof detection [18]. CNN is one of the deep learning architectures that contains multiple convolutional layers to obtain different feature representation from the data. For instance, representation of edges and moles or freckles can be obtained from face image with the first layer and second layer, respectively. In addition to this, BSIF is a feature extraction method which is based on texture of the image data. In that method, firstly, a binary code is computed for each pixel and histogram of pixels' binary code is used for feature representation of image data [19].

1   Istanbul Atlas University, Engineering and Natural Sciences Faculty, Computer Engineering Department, Istanbul, TURKEY.
2   Cyprus Science University, Engineering Faculty, Software Engineering Department, Kyrenia, North Cyprus.
3   Eastern Mediterranean University, Engineering Faculty, Computer Engineering Department, Famagusta, North Cyprus.

Our study employs CNN and BSIF that have been implemented separately for ear anti-spoofing method. Further, they have been fused in the decision-level module of the system. In order to measure the performance of the proposed method, the system is evaluated by using 6 ear databases which are publicly available.

In the remaining part of this study, Section 2 discusses the literature review and Section 3 mentions the characteristics of the system. applied technique is detailed in Section 4. The results of the evaluation of the system are analyzed in Section 5. In the end, study is summarized in Section 6.

LITERATURE REVIEW

Techniques which are used to counter attacks are based on biometric traits such as face, fingerprint, iris, voice, palmprint, finger vein where their texture, motion, frequency, color, shape, quality or reflectance for descriptor approaches are utilized. On the other hand, classifiers that are based on discrimination, regression, distance metrics or heuristics are used for countering against spoof attacks as well [20] [21] [22] [23] [11] [7].

Binarized Statistical Image Features (BSIF) and Convolutional Neural Network (CNN) methods are basically investigated in this study. In the literature, firstly a novel method that employs CNN for multiple channels that are color, depth, infrared and thermal of face images to detect spoof attacks had been proposed. The proposed multi-channel CNN (MC-CNN) had been applied on their self-created face spoof database. In order to compare their method, they have implemented baseline methods on their database. As a result, their method outperforms baseline methods [24]. In the next study, multi-level local binary patterns (MLBP) and CNN methods are employed to obtain different features from face images and resulted with two feature vectors. Further, these vectors are combined to obtain hybrid features. Finally, support vector machine (SVM) is employed by using hybrid features for classification of face images as real or fake [25].

Afterwards, a novel method which employs additive operator splitting (AOS) to detect edges from a face image and specialized CNN architecture to extract features from diffused input images is proposed to resist spoof attacks [26]. Moreover, a face anti-spoofing method that uses deep CNN model where local binary patterns (LBP) method is integrated into its first layer is proposed in [27]. Their proposed method LBPnet is evaluated by conducting experiments on NUAA database. As a result, their proposed method outperforms state-of-the-art methods. Another study has been implemented by using two feature descriptors that are namely multiscale dynamic binarized statistical image features (MBSIF-TOP) and multiscale dynamic local phase quantization (MLPQ-TOP) to find out counterfeit against face identification approach [28]. In order to detect fake face images, discriminative subspace is developed by using spectral regression based kernel discriminant analysis (SR-KDA). Two feature representations which are obtained from MBSIF-

TOP and MLPQ-TOP feature extractors are combined by using SR-KDA method. Consequently, authors have improved the performance of the method by fusing MBSIF-TOP and MLPQ-TOP.

Additionally, there is a method which is developed to detect liveness of a fingerprint biometric trait. In that method, deep learning based deep belief network (DBN) is employed to learn distinctive features from real or fake fingerprint images. They placed the restricted Boltzmann machine (RBM) at each layer in DBN [29]. Next, authors have proposed a spoof detector for fingerprint identification systems. In that proposition, MobileNet-v1 CNN model which is fed with features that are obtained from centered and aligned local patches of minutiae points of fingerprint images is employed. Authors have contributions for improvement of the performance of fingerprint anti-spoofing systems [30]. Beside, BSIF that is texture-based feature extractor method is applied to ensure that fingerprint recognition systems are not accessible by attackers [31] [32].

Furthermore, researchers have focused on identification systems that are based on iris biometric trait to protect against spoofing attacks. Firstly, features are extracted by employing CNN from both local, global and entire iris region images. The obtained feature representation vectors are fused by using feature-level-fusion and score-level-fusion. Next, instead of fully-connected layer of CNN architecture, support vector machine (SVM) is employed for classification part [33]. Additionally, BSIF and CNN-based methods are combined to find robust solution to iris presentation attack detection problem [34]. In that study, different representations are obtained by employing BSIF and these representations are fed to CNN model to obtain multiple classifier results. In order to find the most important and the most prominent view points, meta-analysis algorithm is applied. Finally, Random Forest fusion is used to have a final decision.

On the other hand, CNN-based method is applied for finger vein based presentation attack detection algorithms. In the study [35], they have designed FPNet which is based on CNN model to detect fake finger vein images. The FPNet is conducted on publicly available databases to show its effectiveness.

Moreover, BSIF is implemented for providing a robust method to protect palmprint identification systems. In this context, BSIF is used to learn features from palmprint images and Image Quality Assessment (IQA) methods are used to evaluate the palmprint images in terms of blurrines, deformations and color. Finally, authors employed SVM for classification part [36]. On the other hand, researchers use CNN model in the solution of voice presentation attack detection problem as well [37].

In our previous studies [38] [13], ear anti-spoofing method is proposed with different approaches. In the first study [38], full-reference and no-reference Image Quality Assessment (IQM) metrics had been applied for feature extraction from

ear images. The matching scores had been combined by implementing decision-level fusion technique. In the second study [13], the same IQM metrics are used for feature extraction step. However, score-level fusion (SLF) technique had been employed to combine matching scores. Afterwards, decision-level fusion (DLF) had been applied to obtain a final decision. The experiments had been conducted on AMI and UBEAR ear databases in both studies.

BACKGROUND

In this study, texture-based BSIF method and deep learning based CNN method are utilized to discover powerful and well planned technique in order to find out attacks to ear identification systems. The characteristics of the techniques are clarified as follows:

BINARIZED STATISTICAL IMAGE FEATURES

Binarized Statistical Image Features is a texture-based method for local image description. This method is a variant of Local Binary Patterns (LBP) and Local Phase Quantization (LPQ) approaches. The methodology of this approach is to compute a binary code string for each pixel of an input image by using a learnt filter which is obtained by applying Independent Component Analysis to natural image patches. The linear filter is applied to the image patch that has same size with the filter. The filter response S_i is obtained for each pixel as described in the following formula.

$$S_i = \sum_{u,v} W_i(u,v) * X(u,v) \tag{1}$$

where $W(u,v)$ and $X(u,v)$ represent the linear filter and the image patch, respectively. If the obtained $S_i > 0$, the binary code for that pixel b_i is determined as 1 and 0 otherwise. The contribution of BSIF is to use the statistics of natural image patches to obtain good representation of features [19]. BSIF algorithm is implemented for the applications of identification of a person based on biometric trait and spoof detection systems [28] [39] [40] [41].

CONVOLUTIONAL NEURAL NETWORK

The CNN model is one of the mostly implemented Neural Network (NN) architectures. Its structure exists from hierarchical layers. The reduction of the number of parameters in the CNN model enhances the performance. Additionally, pre-processing is less in the CNN model. CNN model is applied for many research areas such as document recognition, digit classification, text classification, image classification, recommender system, biometric-based recognition system, medical imaging system, speech recognition, natural language processing, spoof detection, gender classification, plant disease and pest detection and character recognition [18] [42] [43] [44] [45] [46] [47] [48] [49] [50]. The general concept of the CNN model exists from multiple layers which are two types namely, convolutional layers and subsampling (pooling) layers.

The main components of conventional Convolutional Neural Network (CNN) can be explained as follows. Firstly, the input data is convolved with different filters by sliding filter over each point of input data to learn features from it. The number of feature maps is equal to the number of filters applied to input data. Multiple convolutional layers are necessary to learn high-level features which represent the input data. In order to detect non-linear features, usage of activation function is inevitable in the convolutional layer. The activation function is applied to all feature values of

obtained feature maps. Some of the activation functions are tanh, sigmoid and ReLU. Further, reduction of dimension of feature maps is achieved by pooling layer which is necessary to provide simpler computation and extract most prominent feature. The mostly used types of pooling layers are average pooling and max pooling. After that, obtained feature maps are combined to get a common feature vector to be fed to fully-connected layer for classification. The backpropagation algorithm is applied in the process of fully-connected layer. The widely used classification technique is Softmax in the CNN model.

PROPOSED METHOD

Our proposed method employs both CNN-based deep learning and texture-based BSIF method to implement a strong way to detect fake ear biometrics. General schema of our suggested technique is shown in Figure 1. The major steps of BSIF part of the algorithm are as follows:

Step 1: Enhance all ear images in train and test sets with histogram equalization and normalize them with mean variance normalization. Next, resize them to 256x256.

Step 2: Convert all ear images into grayscale.

Step 3: Divide each image into 8x8 blocks.

Step 4: Compute BSIF code for each block with 7x7 filter. These operations are applied to all train and test ear images separately.

Step 5: After feature extraction process, train and test sets are shuffled to obtain a realistic model.

Step 6: Nearest Neighbor (NN) method is employed for classification part. The test ear image is compared with all train ear images and scores are obtained by calculating Manhattan distance. NN classifier is applied to determine whether ear image is genuine or impostor.

In addition to this, the detailed steps of CNN part are as follows:

Step 1: Resize all ear images to 256x256 and convert them into grayscale.

Step 2: Train and test sets are shuffled to obtain a realistic model.

Step 3: In the CNN model, 5 convolutional layers are employed with 3x3 kernel map and ReLU activation function.

Step 4: Max Pooling technique with 2x2 pooling size is applied to reduce the dimension of feature map.

Step 5: In order to standardize the inputs of each convolutional layer and speed up training process, batch normalization is applied after pooling operation. Last obtained feature maps are flattened into vector form and it is called input layer. Further, dropout which its rate is set to 20% is applied before input layer to prevent the memorization of data.

Step 6: The input layer is fed to hidden layer which contains 128 neurons to construct a fully-connected NN.

Step 7: Lastly, Softmax classifier is employed in the output layer to categorize test capture as genuine or impostor.
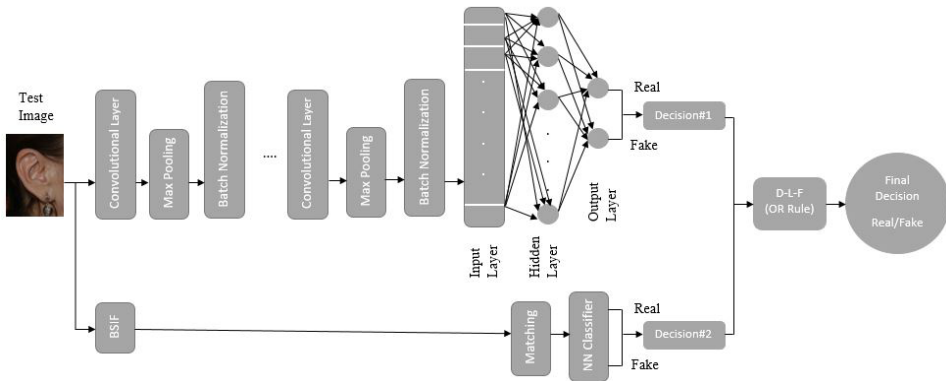
Figure 1. Schema of the proposed ear anti-spoofing method.

CNN and BSIF parts have been implemented separately and two decisions have been obtained. In order to find a common decision, decision-level fusion (DLF) which employs OR method is implemented. The logic of this rule is that if one of the method's decision is genuine, it will lead final decision to be genuine, otherwise it is fake.

EXPERIMENTAL RESULTS

Experiments have been conducted on 6 different ear datasets. The number of real and fake images and original resolution of the images for all datasets are shown in Table 1. We have selected 200 ear images randomly from AMI and UBEAR databases [51] [52]. Further, we have selected 124, 60, 76 and 78 ear samples from IITD and USTB set 1, set 2 and set 3 databases, respectively [53] [54]. Because of absence of spoof database on ear images publicly, we have created spoof databases through printed photo attack on ear images. Therefore, native images have been sent to press with Olivetti d-colour mf223 printer device whose resolution is 1800x600. Additionally, train and test sets have been created with real ear images and corresponding fake ear images.

Table 1. Summary of AMI, UBEAR, IITD and USTB databases.

Database Name	#Real	#Fake	Resolution of image
AMI	200	200	492x702
UBEAR	200	200	1280x960
IITD	124	124	272x204
USTB Set1	60	60	80x150
USTB Set2	76	76	300x400
USTB Set3	78	78	768x576

Performance of the method is measured by applying False Fake Rate (FFR), False Genuine Rate (FGR) and HTER (Half Total Error Rate). FFR is a rate of genuine ear images which are categorized as fake. Whereas, FGR is a rate of impostor ear images which are categorized as real. HTER is the error rate that is calculated as follows:

$$HTER = \frac{FFR + FGR}{2} \qquad (2)$$

Several experiments have been conducted to find the best solution to the problem of ear anti-spoofing. Firstly, BSIF has been implemented on 6 different datasets. As depicted in Table 2 and 3, 1.5, 26.0, 1.0, 8.0, 1.0 and 17.0 error rates have been found out for AMI, UBEAR, IITD, USTB set 1, set 2 and set 3, respectively. Additionally, CNN-based method has been implemented and 0.5, 34.5, 1.0, 4.5, 0.5 and 0.0 error rates have been found out for AMI, UBEAR, IITD, USTB set 1, set 2 and set 3, respectively. CNN-based method achieves minimum error rates on AMI, USTB set 1, set 2 and set 3 datasets. On the other hand, BSIF method achieves minimum error rates on UBEAR dataset. Beside on this, both methods achieve the same performance on IITD dataset. In order to achieve the best performance for all datasets, we have implemented decision-level-fusion (DLF) of BSIF and CNN-based methods. As a result, we have obtained 0.0 error rates for all datasets.

Table 2. Results (%) for each method on AMI, UBEAR and IITD databases.

Method Name	AMI Database			UBEAR Database			IITD Database		
	FFR	FGR	HTER	FFR	FGR	HTER	FFR	FGR	HTER
BSIF	3.0	0.0	1.5	51.0	1.0	26.0	2.0	0.0	1.0
CNN	0.0	1.0	0.5	69.0	0.0	34.5	1.0	1.0	1.0
DLF (BSIF+CNN)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Table 3. Results (%) for each method on USTB set 1, set 2 and set 3 databases.

Method Name	USTB Set 1 Database			USTB Set 2 Database			USTB Set 3 Database		
	FFR	FGR	HTER	FFR	FGR	HTER	FFR	FGR	HTER
BSIF	10.0	6.0	8.0	2.0	0.0	1.0	23.0	11.0	17.0
CNN	3.0	6.0	4.5	1.0	0.0	0.5	0.0	0.0	0.0
DLF (BSIF+CNN)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

COMPARISON WITH THE STATE-OF-THE-ART

In this section, firstly, our proposed method is compared with our previous studies that are focused on ear anti-spoofing problem from different point of view. Those methods employ IQM techniques basically. The first study [38], had been implemented on AMI and UBEAR databases which contain 50 subjects. As shown in Table 4, 8.5 and 15.5 HTER values had been obtained for AMI and UBEAR databases, respectively. The number of subjects for both databases had been incremented to 100 in the second study [13]. In that method which combines SLF and DLF techniques, 1.0 and 10.0 HTER values had been obtained for AMI and UBEAR databases, respectively. Consequently, as it is observed in Table 4, the second study outperforms the first study. Currently, our proposed method is conducted on 6 different ear databases which makes it different from the existing previous experimental studies and instead of using IQM, we employed texture-based and deep learning-based methods in the proposed method. Furthermore, the obtained results that are shown in Table 4 present that our currently proposed method is superior compared to the existing previous studies.

Table 4. Comparison with existing ear anti-spoofing methods.

Reference	Methods Used	Biometric Trait	Databases Used	#Subject	Size of Image	HTER
[38]	IQM (DLF)	Ear	AMI UBEAR	50 50	256x256	8.5 15.5
[13]	IQM(SLF+DLF)	Ear	AMI UBEAR	100 100	256x256	1.0 10.0
Proposed Method	BSIF+CNN (DLF)	Ear	AMI UBEAR IITD USTB Set 1 USTB Set 2 USTB Set 3	100 100 124 60 76 78	256x256	0.0 0.0 0.0 0.0 0.0 0.0

Additionally, CNN-based methods which are performed on biometric traits which are iris, face and fingerprint are compared with our proposed method since there are only two ear anti-spoofing studies in the literature [13] [38]. The summary of the comparison is shown in Table 5. The first study [33] proposes an method for iris biometric trait that is based on CNN + SVM. In that study, 0.016 and 0.292 HTER values are found out for Warsaw-2017 and NDCLD-2015 databases, respectively. In the second study [34], BSIF and CNN are fused for iris presentation attack detection on Notre Dame, Warsaw, Clarkson and IITD+WVU databases and 3.28, 0.68, 9.45 and 14.92 error rates are obtained. In the next study [24], a multi-channel CNN-based method is proposed and conducted on WMCA face spoof database.

Table 5. Comparison with the available methods which use CNN-based deep learning methods on various biometric modalities.

Reference	Year	Methods	Biometric Trait	Databases	#Subject	Image Size	#Epoch	HTER
[33]	2018	CNN+SVM	Iris	Warsaw-2017 NDCLD-2015	- -	- -	-	0.016 0.292
[34]	2018	BSIF+CNN (ensemble of multi view classifiers)	Iris	Notre Dame Warsaw Clarkson IITD+WVU	- - - -	260x260	-	3.28 0.68 9.45 14.92
[24]	2020	MC-CNN	Face	WMCA	72	128x128	-	0.3
[26]	2017	CNN	Face	Replay Attack NUAA	22 50	- 64x64	-	10.0 0.98
[55]	2018	CNN	Face	CASIA-FASD Replay-Attack	50 50	96x96	500	19.12 8.39
[56]	2018	CNN	Face	CASIA-FASD Replay-Attack MSU Rose-Youtu	50 50 35 20	128x128	50	1.4 1.2 0.0 7.0
[57]	2019	CNN	Fingerprint	LivDet2013 LivDet2011	- -	200x200	100	3.7 6.45
[58]	2020	Patch-based deep learning machine (DRBM+DBM Model)	Fingerprint	LivDet2015 (CrossMatch) LivDet2013 (Biometrika) LivDet2013 (ItalData)	51 - -	640x480 315x372 640x480	500	6.44 3.5 2.80
Proposed Method	2020	BSIF+CNN (decision-level fusion)	Ear	AMI UBEAR IITD USTB Set 1 USTB Set 2 USTB Set 3	100 100 124 60 76 78	256x256	250	0.0 0.0 0.0 0.0 0.0 0.0

In that study, HTER value is 0.3. On the other hand, [26] proposed a CNN-based spoof detection algorithm which is conducted on Replay attack and NUAA databases and HTER values of 10.0 and 0.98 are obtained, respectively. Further, CNN-based study [55] on face biometric trait is conducted on CASIA-FASD and Replay attack databases and 19.12 and 8.39 HTER values are obtained, respectively. Additionally, a face anti-spoofing method is proposed by using CNN model [56] and conducted on CASIA-FASD, Replay attack, MSU and Rose-Youtu databases. In that method, HTER values are 1.4, 1.2, 0.0 and 7.0, respectively. Moreover, a CNN-based method [57] is implemented for fake fingerprint detection algorithm. It is conducted on LivDet 2013 and 2011 databases and 3.7 and 6.45 HTER values are obtained. Another fingerprint spoofing study is conducted using patch based deep learning machine with Discriminative Restricted Boltzman Machines (DRBM) and Deep Boltzman Machines (DBM) called DRBM+DBM model [58]. The model obtains HTER values of 6.44, 3.50 and 2.80 on LiveDet2015 (CrossMatch), LiveDet2013 (Biometrika) and LiveDet2013 (ItalData) datasets, respectively.

Finally, our proposed method achieves better results compared to the available anti-spoofing methods with all HTER values equal to 0.0 for all datasets.

CONCLUSION

In this work, a powerful method is proposed by combining deep learning-based CNN and texture-based BSIF methods for ear identification systems security. CNN and BSIF are implemented separately for feature extraction and representation of ear images. In order to show the effectiveness and robustness of the proposed method, 6 distinct ear databases which are AMI, UBEAR, IITD, USTB set1, USTB set2 and USTB set3 are used for the experiments. Combination of BSIF and CNN provides zero half total error rates (HTER) on all aforementioned datasets. Although CNN method and the fusion of CNN and BSIF methods need high computation time, the performance of the proposed method can be considered as the main advantage since zero half total error rates are achieved on 6 ear datasets. In the biometric community, ear anti-spoofing is not studied with CNN and BSIF methods. Consequently, our proposed method has a contribution on the improvement of ear anti-spoofing applications in the biometric community. Further, our proposed method is compared with the available methods which are implemented for face, iris and fingerprint biometric traits by using CNN method and better results are achieved compared to the results of the other biometric anti-spoofing methods.

REFERENCES

[1]   R. D. Rakshit and D. R. Kisku, "Face Spoofing and Counter-Spoofing: A Survey of State-of-the-art Algorithms," Transactions on Machine Learning and Artificial Intelligence, vol. 5, no. 2, pp. 31-73, 2017.

[2]   J. Galbally, S. Marcel and G. Fierrez, "Biometric antispoofing methods: A survey in face recognition," IEEE Access, vol. 2, pp. 1530-1552, 2014.

[3]   S. R. Arashloo, J. Kittler and W. Christmas, "An anomaly detection approach to face spoofing detection: A new formulation and evaluation protocol," IEEE Access, vol. 5, pp. 13868-13882, 2017.

[4]   R. Ramachandra and C. Busch, "Presentation attack detection methods for face recognition systems: A comprehensive survey," ACM Computing Surveys, vol. 50, no. 1, pp. 1-37, 2017.

[5]   O. V. Komogortsev, A. Karpov and C. D. Holland, "Attack of mechanical replicas: Liveness detection with eye movements," IEEE Transactions on Information Forensics and Security, vol. 10, no. 4, pp. 716-725, 2015.

[6]   P. Tome and S. Marcel, "On the vulnerability of palm vein recognition to spoofing attacks," in International Conference on Biometrics (ICB), Phuket, 2015.

[7]   X. Qiu, W. Kang, S. Tian, W. Jia and Z. Huang, "Finger vein presentation attack detection using total variation decomposition," IEEE Transactions on Information Forensics and Security, vol. 13, no. 2, pp. 465-477, 2017.

[8]   B. A. C. S. N. P. S. P. L. Toosi A, A. Toosi, A. Bottino, S. Cumani, P. Negri and P. L. Sottile, "Feature Fusion for Fingerprint Liveness Detection: a Comparative Study," IEEE Access, vol. 5, pp. 23695-23709, 2017.

[9]   S. B. Nikam and S. Agarwal, "Ridgelet-based fake fingerprint detection," Neurocomputing, vol. 72, no. 10-12, pp. 2491-2506, 2019.

[10]   S. Bhilare, V. Kanhangad and N. Chaudhari, "A study on vulnerability and presentation attack detection in palmprint verification system," Pattern Analysis and Applications, vol. 21, no. 3, pp. 769-782, 2018.

[11]   M. Farmanbar and Ö. Toygar, "Spoof detection on face and palmprint biometrics," Signal, Image and Video Processing, vol. 11, no. 7, pp. 1253-1260, 2017.

[12]   B. T. Balamurali, K. E. Lin, S. Lui, J. M. Chen and D. Herremans, "Toward Robust Audio Spoofing Detection: A Detailed Comparison of Traditional and Learned Features," IEEE Access, vol. 7, pp. 84229-84241, 2019.

[13]   İ. Toprak and Ö. Toygar, "Ear anti-spoofing against print attacks using three-level fusion of image quality measures," Signal, Image and Video Processing, vol. 14, no. 2, pp. 417-424, 2020.

[14]   Ž. Emeršič, V. Štruc and P. Peer, "Ear recognition: More than a survey," Neurocomputing, vol. 255, pp. 26-39, 2017.

[15]   E. Alqaralleh and Ö. Toygar, "Ear recognition based on fusion of ear and tragus under different challenges," International Journal of Pattern Recognition and Artificial Intelligence, vol. 32, no. 9, p. 1856009, 2018.

[16]   Ö. Toygar, E. Alqaralleh and A. Alqaralleh, "Symmetric ear and profile face fusion for identical twins and non-twins recognition," Signal, Image and Video Processing, vol. 12, no. 6, pp. 1157-1164, 2018.

[17]   Ö. Toygar, E. Alqaralleh and A. Alqaralleh, "On the use of ear and profile faces for distinguishing identical twins and nontwins," Expert Systems, vol. 37, no. 1, p. e12389, 2020.

[18]   Y. Lecun, Y. Bengio and G. Hinton, "Deep learning," nature, vol. 521, pp. 436-444, 2015.

[19]   J. Kannala and E. Rahtu, "Bsif: Binarized statistical image features," in 21st International Conference on Pattern Recognition (ICPR2012), Tsukuba, 2012.

[20]   L. Souza, L. Oliveira, M. Pamplona and J. Papa, "How far did we get in face spoofing detection?," Engineering Applications of Artificial Intelligence, vol. 72, pp. 368-381, 2018.

[21] E. Marasco and A. Ross, "A survey on antispoofing schemes for fingerprint recognition systems," ACM Computing Surveys, vol. 47, no. 2, p. 28, 2015.

[22] A. Czajka and K. W. Bowyer, "Presentation attack detection for iris recognition: An assessment of the state-of-the-art," ACM Computing Surveys, vol. 51, no. 4, p. 86, 2018.

[23] H. Muckenhirn, M. Magimai-Doss and S. Marcel, "End-to-end convolutional neural network-based voice presentation attack detection," in IEEE International Joint Conference on Biometrics (IJCB), Denver, CO, 2017.

[24] A. George, Z. Mostaani, D. Geissenbuhler, O. Nikisins, A. Anjos and S. Marcel, "Biometric Face Presentation Attack Detection with Multi-Channel Convolutional Neural Network," IEEE Transactions on Information Forensics and Security, vol. 15, pp. 42-55, 2020.

[25] D. T. Nguyen, T. D. Baek and K. R. Park, "Combining deep and handcrafted image features for presentation attack detection in face recognition systems using visible-light camera sensors," Sensors, vol. 18, no. 3, p. 699, 2018.

[26] A. Alotaibi and A. Mahmood, "Deep face liveness detection based on nonlinear diffusion using convolution neural network," Signal, Image and Video Processing, vol. 11, no. 4, pp. 713-720, 2017.

[27] G. B. De Souza, D. F. da Silva Santos, R. G. Pires, A. N. Marana and J. P. Papa, "Deep texture features for robust face spoofing detection," IEEE Transactions on Circuits and Systems II: Express Briefs, vol. 64, no. 12, pp. 1397-1401, 2017.

[28] S. R. Arashloo, J. Kittler and W. Christmas, "Face spoofing detection based on multiple descriptor fusion using multiscale dynamic binarized statistical image features," IEEE Transactions on Information Forensics and Security, vol. 10, no. 11, pp. 2396-2407, 2015.

[29] S. Kim, B. Park, B. S. Song and S. Yang, "Deep belief network based statistical feature learning for fingerprint liveness detection," Pattern Recognition Letters, vol. 7, pp. 58-65, 2016.

[30] T. Chugh, K. Cao and A. K. Jain, "Fingerprint spoof buster: Use of minutiae-centered patches," IEEE Transactions on Information Forensics and Security, vol. 13, no. 9, pp. 2190-2202, 2018.

[31] L. Ghiani, A. Hadid, G. L. Marcialis and F. Roli, "Fingerprint liveness detection using binarized statistical image features," in IEEE Sixth International Conference on Biometrics: Theory, Applications and Systems (BTAS), Arlington, VA, 2013.

[32] Q. Li and P. P. K. Chan, "Fingerprint liveness detection based on binarized statistical image feature with sampling from Gaussian distribution," in 2014 International Conference on Wavelet Analysis and Pattern Recognition, Lanzhou, 2014.

[33] D. T. Nguyen, T. D. Pham, Y. W. Lee and K. R. Park, "Deep learning-based enhanced presentation attack detection for iris recognition by combining features from local and global regions based on NIR camera sensor," Sensors, vol. 18, no. 8, p. 2601, 2018.

[34] A. Kuehlkamp, A. Pinto, A. Rocha, K. W. Bowyer and A. Czajka, "Ensemble of multi-view learning classifiers for cross-domain iris presentation attack detection," IEEE Transactions on Information Forensics and Security, vol. 14, no. 6, pp. 1419-1431, 2018.

[35] X. Qiu, S. Tian, W. Kang, W. Jia and Q. Wu, "Finger Vein Presentation Attack Detection Using Convolutional Neural Networks," in Chinese Conference on Biometric Recognition (CCBR), Cham, Springer, 2017, pp. 296-305.

[36] X. Li, W. Bu and X. Wu, "Palmprint liveness detection by combining binarized statistical image features and image quality assessment," in Chinese Conference on Biometric Recognition (CCBR), Cham, Springer, 2015, pp. 275-283.

[37] H. Muckenhirn, P. Korshunov, M. Magimai-Doss and S. Marcel, "Long-term spectral statistics for voice presentation attack detection," IEEE/ACM Transactions on Audio, Speech and Language Processing (TASLP), vol. 25, no. 11, pp. 2098-2111, 2017.

[38] İ. Toprak and Ö. Toygar, "Fusion of Full-Reference and No-Reference Anti-Spoofing Techniques for Ear Biometrics under Print Attacks," in International Conference on Advanced Technologies, Computer Engineering and Science (ICATCES), Karabuk, 2018.

[39] A. Ouamane, A. Benakcha, M. Belahcene and A. Taleb-Ahmed, "Multimodal depth and intensity face verification approach using LBP, SLF, BSIF, and LPQ local features fusion," Pattern Recognition and Image Analysis, vol. 25, no. 4, pp. 603-620, 2015.

[40] R. Raghavendra and C. Busch, "Robust scheme for iris presentation attack detection using multiscale binarized statistical image features," IEEE Transactions on Information Forensics and Security, vol. 10, no. 4, pp. 703-715, 2015.

[41] J. S. Doyle and K. W. Bowyer, "Robust Detection of Textured Contact Lenses in Iris Recognition Using BSIF," IEEE Access, vol. 3, pp. 1672-1683, 2015.

[42] J. Gu, Z. Wang, J. Kuen, L. Ma, A. Shahroudy, B. Shuai, T. Liu, X. Wang, G. Wang, J. Cai and T. Chen, "Recent advances in convolutional neural networks," Pattern Recognition, vol. 77, pp. 354-377, 2018.

[43] W. Liu, Z. Wang, X. Liu, N. Zeng, Y. Liu and F. E. Alsaadi, "A survey of deep neural network architectures and their applications," Neurocomputing, vol. 234, pp. 11-26, 2017.

[44] X. Chen and X. Lin, "Big Data Deep Learning: Challenges and Perspectives," IEEE Access, vol. 2, pp. 514-525, 2014.

[45] D. Ramachandram and G. W. Taylor, "Deep multimodal learning: A survey on recent advances and trends," IEEE Signal Processing Magazine, vol. 34, no. 6, pp. 96-108, 2017.

[46] G. Wang, "A Perspective on Deep Imaging," IEEE Access, vol. 4, pp. 8914-8924, 2016.

[47] S. A. Radzi, M. K. Hani and R. Bakhteri, "Finger-vein biometric identification using convolutional neural network," Turkish Journal of Electrical Engineering & Computer Sciences, vol. 24, no. 3, pp. 1863-1878, 2016.

[48] S. S. Liew, M. K. Hani, S. A. Radzi and R. Bakhteri, "Gender classification: a convolutional neural network approach," Turkish Journal of Electrical Engineering & Computer Sciences, vol. 24, no. 3, pp. 1248-1264, 2016.

[49] M. Türkoğlu and D. Hanbay, "Plant disease and pest detection using deep learning-based features," Turkish Journal of Electrical Engineering & Computer Sciences, vol. 27, no. 3, pp. 1636-1651, 2019.

[50] M. A. Rizvi, K. Deb, M. I. Khan, M. M. S. Kowsar and T. Khanam, "A comparative study on handwritten Bangla character recognition," Turkish Journal of Electrical Engineering & Computer Sciences, vol. 27, no. 4, pp. 3195-3207, 2019.

[51] E. Gonzalez, L. Alvarez and L. Mazorra, "AMI Ear Database," [Online]. Available: http://www.ctim.es/research_works/ami_ear_database. [Accessed 10 January 2018].

[52] R. Raposo, E. Hoyle, A. Peixinho and H. Proença, "Ubear: A dataset of ear images captured on-the-move in uncontrolled conditions," in IEEE Workshop on Computational Intelligence in Biometrics and Identity Management (CIBIM), 2011.

[53] A. Kumar and C. Wu, "Automated human identification using ear imaging," Pattern Recognition, vol. 45, no. 3, pp. 956-968, 2012.

[54] Z. Mu and L. Yuan, "Ear Recognition Laboratory at USTB," [Online]. Available: http://www1.ustb.edu.cn/resb/en/index.htm. [Accessed 15 December 2018].

[55] Y. A. U. Rehman, L. M. Po and M. Liu, "LiveNet: Improving features generalization for face liveness detection using convolution neural networks," Expert Systems with Applications, vol. 108, pp. 159-169, 2018.

[56] H. Li, P. He, S. Wang, A. Rocha, X. Jiang and A. C. Kot, "Learning generalized deep feature representation for face anti-spoofing," IEEE Transactions on Information Forensics and Security, vol. 13, no. 10, pp. 2639-2652, 2018.

[57] C. Yuan, Z. Xia, L. Jiang, Y. Cao, Q. M. J. Wu and X. Sun, "Fingerprint Liveness Detection Using an Improved CNN With Image Scale Equalization," IEEE Access, vol. 7, pp. 26953-26966, 2019.

[58] D. M. Uliyan, S. Sadeghi and H. A. Jalab, "Anti-spoofing method for fingerprint recognition using patch based deep learning machine," Engineering Science and Technology an International Journal- Jestech, vol. 23, pp. 264-273, 2020.

REDUCED DIFFERENTIAL TRANSFORM METHOD WITH FIXED GRID SIZE FOR SOLVING GAOUSAT PROBLEM

Sema SERVİ [1]

INTRODUCTION

Problems encountered in many fields of science through applied mathematics can be solved with mathematical models. The Goursat problem is a hyperbolic partial differential equation that appears in the analysis and calculation of many science and engineering problems. Many other problems such as stationary mass equations, Dirac and Maxwell fields, stochastic systems are solved with the Goursat problem. The solution is calculated serially and converges to the exact solution. In this study, the Goursat problem is solved with the Fixed Grid Dimension Reduced Differential Transformation Method (RDTM vs. FGS), different from many problems solved so far, and compared with methods such as, Adomian decomposition method (ADM), the variational iteration method (VIM) in the literature. Thus, the success, reliability and effectiveness of the method have been demonstrated.

The mathematical models we create to solve problems in many scientific fields usually consist of partial differential equations. Partial differential equations can appear as linear and non-linear equations. Solving these equations analytically is not possible in most cases. It is also highly time consuming and difficulting when it can be resolved. For this reason, many approximate solution methods are used to solve partial differential equations. Variational iteration method [13,16-18,21], modified variational iteration algorithm-II [18], the homotopy analysis method [19,20], homotopy perturbation method [20], the differential transform method (DTM) [14,15,21], the Adomian's decomposition method (ADM) [8,10-12,31], reduced differential transform method (RDTM) [1-4,5,6,7,8,9], Spline collacation medhod [22], Finite Difference Method [23,24], Laplace transform method [25-27] and so forth.

RDTM with FGS is a method created by the development of RDTM. With this method, firstly, partial differentiable differential equations are converted into algebraic equations. Then, the interval in which the equation is solved is divided into equal parts and approximate results are obtained with the fixed grid size algorithm and simple algebraic operations. Fixed grid size algorithms were applied with DTM for the approximate solution of linear and nonlinear initial value problems

---

1   Selcuk University, Faculty of Technology, Computer Engineering, Konya, TURKEY.

in previous studies [2-3] and efficient results were obtained. In this study, RDTM with FGS was applied to the Goursat problem, which is encountered in many areas in the literature. Many problems such as generalized spin zero rest mass equations on Minkowski space-time, between the black hole event horizon and the Cauchy horizon, Dirac fields inside a Reissner-Nordström-like black hole, an optimal control problem for the stochastic system, are solved by the Goursat problem. General form of the Goursat equation is given by the following equation:

$$u(x,t) = f(x,t,u_x,u_t), 0 \leq x \leq a, 0 \leq t \leq b \tag{1.1}$$

subject to initial and boundary conditions,

$$u(x,0) = g(x), \; u(0,t) = h(t), \; g(0) = h(0) = u(0,0) \tag{1.2}$$

RDTM WITH FIXED GRID SIZE SOLUTION

The principal descriptions of RDTM are given in the references [1-6,28-30]:

Definition 2.1. If the function $u(x,t)$ is analytic and differentiated continuously with respect to time t and space x in the domain of interest, the equation where the t-dimensional spectrum function $U_k(x)$ is the transformed function is like this:

$$U_k(x) = \frac{1}{k!}\left[\frac{\partial^k}{\partial t^k} u(x,t)\right]_{t=0} \tag{2.1}$$

The differential inverse transform of $U_k(x)$ is defined as in the following:

$$u(x,t) = \sum_{k=0}^{\infty} U_k(x)t^k \tag{2.2}$$

When equations of (2.1) and (2.2) are combined, we write,

$$u(x,t) = \sum_{k=0}^{\infty} \frac{1}{k!}\left[\frac{\partial^k}{\partial t^k} u(x,t)\right]_{t=0} t^k \tag{2.3}$$

As it can be understood from the definitions given, the concept of reduced differential transformation is created by using the power series expansion.

According to the RDTM and Table 1, we can derive the iteration formulas.:

Table 1. Reduced differential transformation

Functional	Transformed Form
$u(x,t)$	$U_k(x) = \dfrac{1}{k!}\left[\dfrac{\partial^k}{\partial t^k}u(x,t)\right]_{t=0}$
$w(x,t) = u(x,t) \pm v(x,t)$	$W_k(x) = U_k(x) \pm V_k(x)$
$w(x,t) = \alpha u(x,t)$	$W_k(x) = \alpha U_k(x)$ (α is a constant)
$w(x,y) = x^m t^n$	$W_k(x) = x^m \delta(k-n)$
$w(x,y) = x^m t^n u(x,t)$	$W_k(x) = x^m U(k-n)$
$w(x,t) = u(x,t)v(x,t)$	$W_k(x) = \sum_{r=0}^{k} V_r(x) U_{k-r}(x) = \sum_{r=0}^{k} U_r(x) V_{k-r}(x)$
$w(x,t) = \dfrac{\partial^r}{\partial t^r}u(x,t)$	$W_k(x) = (k+1)\dots(k+r)U_{k+r}(x) = \dfrac{(k+r)!}{k!}U_{k+r}(x)$
$w(x,t) = \dfrac{\partial}{\partial x}u(x,t)$	$W_k(x) = \dfrac{\partial}{\partial x}U_k(x)$

With the grid points $\{t_0, t_1, t_2, \dots, t_N\}$ pointed out, the interval is divided into N evenly spaced subintervals. Our goal in this article is to investigate the solution of (1.2) for the interval [0,T].

$t_0 = 0, t_N = T$ as in the formulation,

$t_i = t_0 + ih$, for each i=0,1, 2,…,N and $h = \dfrac{T}{N}$

(2.4)

The approximate solution function $u(x,t)$ is designated as $u^0(x,t)$ in the first interval, and RDTM techniques are then used on the problem to get the nth order Taylor polynomials at point $t_0 = 0$. Equation (2.5) is produced as a result.

$u^0(x,t) = U_0^0(x,t_0) + U_1^0(x,t_0)(t-t_0) +$
$U_2^0(x,t_0)(t-t_0)^2 + \dots + U_n^0(x,t_0)(t-t_0)^n$

(2.5)

From the initial condition (1.2),

$u(x,t_0) = U_0^0(x,t_0)$

(2.6)

From here, we can determine from (2.5) the approximation of the u function at t_1 as follows.

$u(x,t_1) \approx u^0(x,t_1)$
$= U_0^0(x,t_0) + U_1^0(x,t_0)(t_1-t_0) + U_2^0(x,t_0)(t_1-t_0)^2 +$
$\dots + U_n^0(x,t_0)(t_1-t_0)^n$
$= U_0^0 + U_1^0 h + U_2^0 h^2 + \dots + U_n^0 h^n$
$= \sum_{j=0}^{n} U_j^0 h^j$

(2.7)

The following equation can be constructed because the initial value of the second sub-domain at the instant t_1 is equal to $u^0(x,t)$:

$$u^1(x,t) = U_0^1(x,t_1) + U_1^1(x,t_1)(t-t_1) + U_2^1(x,t_1)(t-t_1)^2 +$$
$$\ldots + U_n^1(x,t_1)(t-t_1)^n$$

$$u^1(x,t_1) = u^0(x,t_1) = U_0^1(x,t_1)$$

(2.8)

In a same way, $u^2(x,t_2)$ can be given as,

$$u(x,t_2) \approx u^1(x,t_2)$$
$$= U_0^1(x,t_1) + U_1^1(x,t_1)(t_2-t_1) + U_2^1(x,t_1)(t_2-t_1)^2 +$$
$$\ldots + U_n^1(x,t_1)(t_2-t_1)^n$$
$$= U_0^1 + U_1^1 h + U_2^1 h^2 + \ldots + U_n^1 h^n$$
$$= \sum_{j=0}^{n} U_j^1 h^j$$

(2.9)

Following the similar processes, we can arrive at the following grid point t_{i+1} solution for $u^i(x,t)$:

$$u(x,t_{i+1}) \approx u^i(x,t_{i+1})$$
$$= U_0^i(x,t_i) + U_1^i(x,t_i)(t_{i+1}-t_i) + U_2^i(x,t_i)(t_{i+1}-t_i)^2 +$$
$$\ldots + U_n^i(x,t_i)(t_{i+1}-t_i)^n$$
$$= U_0^i + U_1^i h + U_2^i h^2 + \ldots + U_n^i h^n$$
$$= \sum_{j=0}^{n} U_j^i h^j$$

(2.10)

After calculating the $u^i(x,t)$ values from these equations, the exact solution of $u(x,t)$ is obtained as follows:

$$u(x,t) = \begin{cases} u^0(x,t), & 0 \le x \le 1,\ 0 < t \le t_1 \\ u^1(x,t), & 0 \le x \le 1,\ t_1 \le t \le t_2 \\ \vdots \\ u^{N-1}(x,t), & 0 \le x \le 1,\ t_{N-1} \le t \le t_N \end{cases}$$

It can be understood from here that as we raise the value of N, we get more sensitive results.

APPLICATION

In this section, the method is compared with known exact solution and different approach solution methods to evaluate the efficiency and success of our method. The numerical results obtained are very encouraging.

Example 3.1:

The following equation describes the homogeneous Goursat problem. [8,12,13,31]:

$$u_{xt} = u \tag{3.1}$$

The initial conditions are given by,

$$u(x,0) = e^x, u(0,t) = e^t, u(0,0) = 1 \tag{3.2}$$

and the precise answer to the (3.1) equation is,

$$u(x,t) = e^{(x+t)} \tag{3.3}$$

for N=10, we obtain the t-values as follows:

$$\left\{ \begin{array}{l} t_0 = 0, t_1 = \dfrac{1}{10}, t_2 = \dfrac{2}{10}, t_3 = \dfrac{3}{10}, t_4 = \dfrac{4}{10}, t_5 = \dfrac{5}{10}, \\ t_6 = \dfrac{6}{10}, t_7 = \dfrac{7}{10}, t_8 = \dfrac{8}{10}, t_9 = \dfrac{9}{10}, t_{10} = 1 \end{array} \right\}$$

For the solution the equation, we first get the RDTM transform of (3.1) and (3.2) by the aid of Table 2.1 and we obtain the following equation,

$$(k+1)\frac{d}{dx}U_{k+1}^0 = U_k^0, k \geq 0$$

The initial conditions are given by,

$$U_0^0(x,t_0) = e^x$$

Table 2 below shows the absolute error values at t and x points for the solutions of equation (3.1) with RDTM with FGS, VIM [13], ADM [8,12,13,31]. In Figure 1-2, solution and error graphs of RDTM with FGS, VIM and ADM are given. As seen in Table 2 and Figure 1-2, the best approximation to the exact solution among the obtained results was calculated with RDTM with FGS. This clearly demonstrates the success of the method.

Table 2: Numerical Results of RDTM with FGS solution for N=10

	x	Absolute error (RDTM with FGS)	Absolute error (VIM)	Absolute error (ADM)
	0.2	$1.3152921910.10^{-9}$	$3.1812830635.10^{-9}$	$3.357231486351.10^{-7}$
	0.4	$1.6065015099.10^{-9}$	$3.8856279083.10^{-9}$	0.0000055786804455288
t=0.2	0.6	$1.9621853752.10^{-9}$	$4.7459166443.10^{-9}$	0.0000388392974777046
	0.8	$2.3966186293.10^{-9}$	$5.7966756794.10^{-9}$	0.0001698713543172682
	0.2	$3.7369179924.10^{-9}$	$4.178244355378.10^{-7}$	0.0000055786804455288
t=0.4	0.4	$4.564281943.10^{-9}$	$5.103319179926.10^{-7}$	0.0000451522730760636
	0.6	$5.5748265542.10^{-9}$	$6.23320812213.10^{-7}$	0.0002166980519206854
	0.8	$6.8091085297.10^{-9}$	$7.613257592559.10^{-7}$	0.0007640794831402067

t=0.6	0.2	$6.554951645.10^{-9}$	$7.3288935177001.10^{-6}$	0.0000388392974777046
	0.4	$8.006236019.10^{-9}$	0.0000089515307567811	0.0002166980519206854
	0.6	$9.7788387567.10^{-9}$	0.0000109334243560878	$8.118558878353139.10^{-4}$
	0.8	$1.1943900629.10^{-8}$	0.0000133541146646613	0.0024261245980376810
t=0.8	0.2	$8.8875428900.10^{-9}$	0.0000563957524967322	0.0001698713543172682
	0.4	$1.0855269399.10^{-9}$	0.0000688819276480269	0.0007640794831402067
	0.6	$1.3258655984.10^{-8}$	$8.41325764166893.10^{-5}$	0.0024261245980376810
	0.8	$1.61941589893.10^{-8}$	0.0001027597608864656	0.0064105493177863166
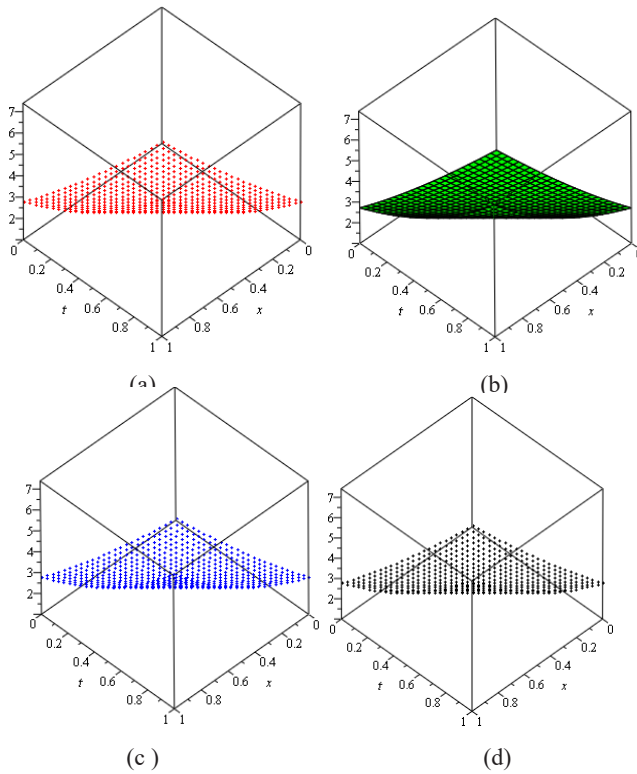
Figure 1. (a): RDTM with FGS solution (b): Exact solution (c): VIM solution (d): ADM solution

(a)

(b)

(c)
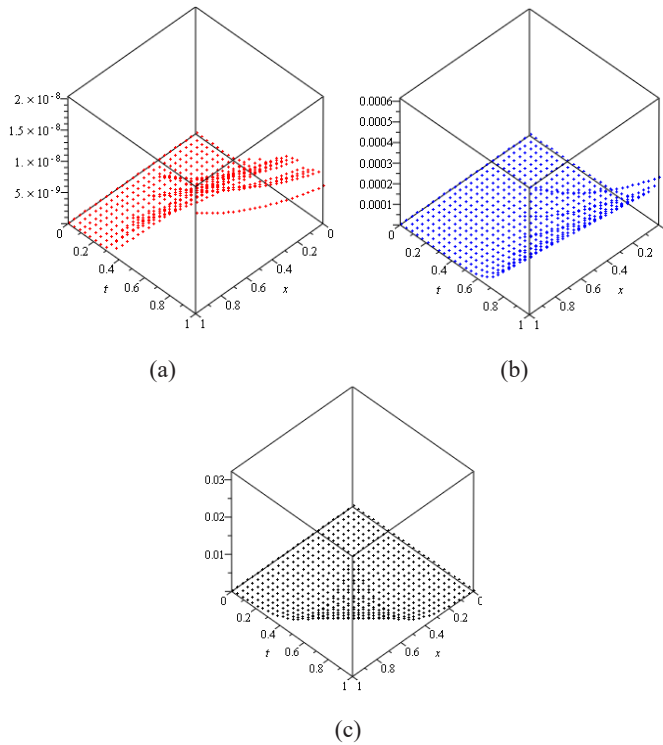
Figure 2. (a): Absolute error for RDTM with FGS solution (b): Absolute error for VIM solution
(c): Absolute error for ADM solution

Example 3.2:

The linear homogeneous Goursat Problem ,is given by the following equation [8,12,13,31],

$$u_{xt} = -2u \tag{3.4}$$

with the given conditions,

$$u(x,0) = e^x , u(0,t) = e^t , u(0,0) = 1 \tag{3.5}$$

and the analytical solution for the (3.1) equation is,

$$u(x,t) = e^{x-2t}$$

Using the transformation equivalents of equation (3.4) from Table 1, the following equation is obtained,

$$(k+1)\frac{dU^0_{k+1}(x)}{dx} = -2U^0_k$$

and initial conditions,

$$U^0_0(x,t_0) = e^x$$

Figure 3's results VIM-Exact and RDTM with FGS-Exact are shown below [5]. Since the results obtained in the solution made with ADM give more errors than VIM and RDTM with FGS, they are not included in the evaluation in this question.

(a)   (b)
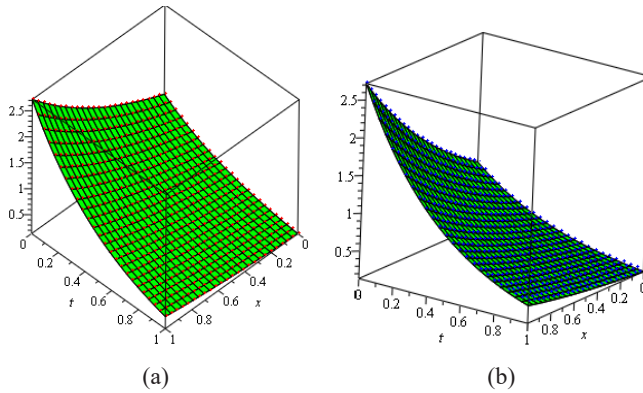
Figure 3. (a): RDTM with FGS -Exact solution (b): VIM-Exact solution

Table 3 below shows the absolute error values at t and x points for the solutions of equation (3.1) with RDTM with FGS, VIM [13].

Table 3: Numerical Results of RDTM with FGS solution for N=10

	x	Absolute error (RDTM with FGS)	Absolute error (VIM)
	0.2	$6.9327778856.10^-$	$3.7804847941912.10^{-7}$
	0.4	$8.4677140313.10^{-9}$	$4.61749455480.10^{-7}$
t=0.2	0.6	$1.03424892731.10^{-8}$	$5.639820585032.10^{-7}$
	0.8	$1.26323449244.10^{-8}$	$6.888492418087.10^{-7}$
	0.2	$8.11090794330.10^{-9}$	0.00004615856701594375
t=0.4	0.4	$9.90668533313.10^{-9}$	0.00005637820106599474
	0.6	$1.210005279013.10^{-8}$	0.00006886049028211463
	0.8	$1.47790378517.10^{-8}$	0.0000841063927588364
	0.2	$8.40617391917.10^{-9}$	0.00075359095218832817
t=0.6	0.4	$1.026732401045.10^{-8}$	0.00092043806752737271
	0.6	$1.254053786531.10^{-8}$	0.00112422559439354970
	0.8	$1.531704753750.10$	0.00137313224178653797
	0.2	$6.88645629875.10^{-9}$	$0.0054032873985 3599792$
t=0.8	0.4	$8.41113671724.10^{-9}$	$0.0065995901317 0395668$
	0.6	$1.027338558570.10^{-8}$	$0.0080607575895 8985118$
	0.8	$1.254794149003.10^{-8}$	$0.0098454315527 8556653$

Figure 4 shows absolute error graphs of RDTM with FGS and VIM solutions.

(a)   (b)
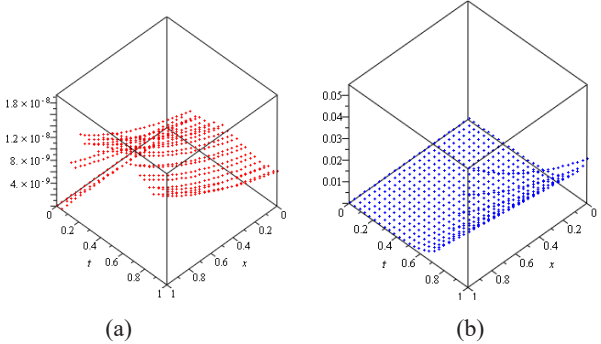
Figure 4. (a): Absolute error for RDTM with FGS solution (b): Absolute error for VIM solution

CONCLUSION

It is extremely important to be able to solve linear and non-linear partial differential equations, which are encountered in solving or modeling many engineering problems. Given that solving partial differential equations is extremely challenging, most of them cannot be solved analytically, so approximate solution methods gain importance. In this study, the Goursat problem, which is encountered in many engineering problems, was solved with RDTM with FGS, a method different from the methods solved in the literature. And the results were also compared with VIM and ADM. In the results obtained, it is seen that the RDTM with FGS method approaches the exact solution more quickly with an analytical approach compared to VIM, ADM, and also gives much more successful results. This success also shows promise in solving linear and non-linear partial differential equations of the RDTM with FGS method, which are also encountered in other problems. In addition, RDTM with FGS is an extremely advantageous method in that it can be programmed very easily and can give very good results with less iteration. In this study, the calculations were made with the help of the Maple program.

REFERENCES

[1] Servı, Sema, and Galip Oturanç. "Reduced Differential Transform Approach Using Fixed Grid Size for Solving Newell–Whitehead–Segel (NWS) Equation." The International Conference on Artificial Intelligence and Applied Mathematics in Engineering. Springer, Cham, 2020.

[2] Jang, Ming-Jyi, Chieh-Li Chen, and Yung-Chin Liy. "On solving the initial-value problems using the differential transformation method." Applied Mathematics and Computation 115.2 (2000): 145-160.

[3] Kurnaz, Aydin, and Galip Oturanç. "The differential transform approximation for the system of ordinary differential equations." International Journal of Computer Mathematics 82.6 (2005): 709-719.

[4] Keskin, Yildiray, and Galip Oturanc. "Reduced differential transform method for partial differential equations." International Journal of Nonlinear Sciences and Numerical Simulation 10.6 (2009): 741-750.

[5] Keskin, Yıldıray, and Galip Oturanç. "Reduced differential transform method for generalized KdV equations." Mathematical and Computational applications 15.3 (2010): 382-393.

[6] Servi, Sema, Yildiray Keskin, and Galip Oturanç. "Reduced differential transform method for improved Boussinesq equation." Proceedıngs of the internatıonal conference on numerical analysis and applied mathematics 2014 (ICNAAM-2014). Vol. 1648. No. 1. AIP Publishing, 2015.

[7] Appadu, Appanah Rao, and Abey Sherif Kelil. "Solution of 3D linearized KdV equation using reduced differential transform method." AIP Conference Proceedings. Vol. 2425. No. 1. AIP Publishing LLC, 2022.

[8] Naseem, Tahir. "Novel techniques for solving Goursat partial differential equations in the linear and nonlinear regime." International Journal of Emerging Multidisciplinaries: Mathematics 1.1 (2022): 17-37.

[9] *Aljahdaly, Noufe H., and Maram A. Alharbi. "On Reduce Differential Transformation Method for Solving Damped Kawahara Equation." Mathematical Problems in Engineering 2022 (2022).*

[10] Al-Mazmumy, Mariam A. "Adomian decomposition method for solving Goursat's problems." Applied Mathematics 2.8 (2011): 975-980.

[11] Deraman, Ros Fadilah, et al. "Adomian decomposition associated with Newton-Cotes formula for solving Goursat problem." AIP Conference Proceedings. Vol. 1522. No. 1. American Institute of Physics, 2013.

[12] Wazwaz, Abdul-Majid. "The decomposition method for approximate solution of the Goursat problem." Applied Mathematics and Computation 69.2-3 (1995): 299-311.

[13] Wazwaz, Abdul-Majid. "The variational iteration method for a reliable treatment of the linear and the nonlinear Goursat problem." Applied Mathematics and Computation 193.2 (2007): 455-462.

[14] Idrees, M., et al. "Exact solution of Goursat problems using differential transform method." Journal of Advanced Research in Scientific Computing 3.3 (2011): 1-13.

[15] Taghvafard, H., and G. H. Erjaee. "Two-dimensional differential transform method for solving linear and non-linear Goursat problem." International Journal for Engineering and Mathematical Sciences 6.2 (2010): 103-106.

[16] Noor, Muhammad Aslam, and Syed Tauseef Mohyud-Din. "Modified variational iteration method for Goursat and Laplace problems." World Applied Sciences Journal 4.4 (2008): 487-498.

[17] Al-Fayadh, Ali, and Dina Saad Faraj. "Laplace Substitution–Variational Iteration Method for Solving Goursat Problems Involving Mixed Partial Derivatives." American Journal of Mathematical and Computer Modelling 4.1 (2019): 16-20.

[18] Mohyud-Din, Syed Tauseef, Muhammad Aslam Noor, and Khalida Inayat Noor. "Solving Goursat and Laplace problems by modified variational iteration method." Nonl. Anal. Forum. Vol. 14. 2009.

[19] Butt, Adnan Saeed, et al. "Effect of viscoelasticity on entropy generation in a porous medium over a stretching plate." World Applied Sciences Journal 17.4 (2012): 516-523.

[20] Yıldırım, Ahmet, and Meryem Odabaşı. "The homotopy perturbation method for solving the linear and the nonlinear Goursat problems." International Journal for Numerical Methods in Biomedical Engineering 27.7 (2011): 1139-1148.

[21] Deraman, Ros Fadilah. A new scheme for solving goursat problem using newton-cotes integration formula. Diss. Universiti Teknologi MARA, 2014.

[22] Agachev, Yurii Romanovich. "Approximate solution of the Goursat problem by the spline-collocation method." Izvestiya Vysshikh Uchebnykh Zavedenii. Matematika 6 (1979): 66-69.

[23] Pandey, Pramod Kumar. "A Finite Difference Method for Numerical Solution of Goursat Problem of Partial Differential Equation." Open Access Libr J 1 (2014): 1-6.

[24] Aziz, A. K., and B. E. Hubbard. "Bounds on the truncation error by finite differences for the Goursat problem." Mathematics of computation 18.85 (1964): 19-35.

[25] Leble, S. B., and A. V. Yurov. "Reduction restrictions of Darboux and Laplace transformations for the Goursat equation." Journal of Mathematical Physics 43.2 (2002): 1095-1105.

[26] Cirnu, Mircea Ion. "Goursat problem for some hyperbolic equations solved by Laplace transform method." Int. J. Modern Sci. Eng. Technol 1 (2014): 9-15.

[27] Al-Fayadh, Ali, and Dina Saad Faraj. "Laplace Substitution–Variational Iteration Method for Solving Goursat Problems Involving Mixed Partial Derivatives." American Journal of Mathematical and Computer Modelling 4.1 (2019): 16-20.

[28] Servi, Sema, Yıldıray Keskin, And Galip Oturanc. "Approximate Solution For Solving The Sine-Gordon Equation By Reduced Differential Transform Method With Fixed Grid Size." 2nd International Conference On Analysis And Its Applications July, 12-15, 2016, KIRSEHIR/TURKEY.

[29] Servi, Sema, Yildiray Keskin, And Galip Oturanç. "Reduced Differential Transform Method With Fixed Grid Size For Solving Telegraph Equations." Book of Abstracts Computational Methods in Applied Mathematics (CMAM-7) July 31-August 6, 2016 University of Jyväskylä, Finland (2016): 129.

[30] Servi, Sema, Yildiray Keskin, And Galip Oturanç. "Numerical Solution Of Wave Equations By Reduced Differential Transform Method With Fixed Grid Size." Book of Abstracts Computational Methods in Applied Mathematics (CMAM-7) July 31-August 6, 2016 University of Jyväskylä, Finland (2016): 128.

[31] Mohmoud, Sharaf, and Mohamed Gubara. "Reduced differential transform method for solving linear and nonlinear goursat problem." Applied Mathematics 7.10 (2016): 1049-1056.

REWARDING WATER ANALYSIS DATA IN PROOF OF STAKE BLOCKCHAIN NETWORK

Batuhan GÖKTAŞ[1,] Nurettin DOĞAN[2]

INTRODUCTION

Blockchain is a common database that uses cryptographic methods to secure data and organizes data groups in a certain order according to time [1]. This database of common records contains data that is open and long-standing and that, among other things, occurs in a particular established group, such as historical transaction data [2]. A distributed node consensus algorithm, cryptography methods, smart contracts consisting of automated script code, distributed infrastructure, computing techniques, etc. are some of the key components of blockchain technology [3]. The above-mentioned components enable blockchain technology to meet the data specifications including the creation, verification, storage, and updating of data and the security of the data [4]. Decentralization is also advantageous for blockchain at the same time. Bitcoin is this technology's most well-known application in the modern world [5]. Its benefit is the elimination of intermediary parties, and docking currency passwords is more practical.

Because data can be written to and saved on the block chain, which is primarily used to store data, the block chain functions as a database. Chain block chains can be divided into three categories: public, private, and union chain strands. Each of these categories has unique properties, although they can partially overlap. Users can select the appropriate blockchain type for their needs. The basic building parts of the block chain are the data layer, network layer, consensus layer, incentive layer group, contract layer, and consensus layer. The developments in block chain technology that best exemplify economic incentive schemes and smart contracts are those that are based on the time chain link structure, distributed node consensus mechanism, and consensus idea. Although not a fundamental component of block chain design, the incentive layer, contract layer, and application layer are generally used to create decentralized applications. Instead, they are needed to build consensus architecture block chains and cannot be deemed complete in their absence [6].

In addition to being directly responsible for protecting their health, people also need to think about the health of the people around them, and they must contribute

1   Batuhan GÖKTAŞ, Department of Computer Engineering, Faculty of Technology, Selçuk University, Alaeddin Keykubat Campus, Konya, 42075, Türkiye, batuhangoktas@hotmail.com , 0000-0001-5675-6913
2   Nurettin DOĞAN, Department of Computer Engineering, Faculty of Technology, Selçuk University, Alaeddin Keykubat Campus, Konya, 42075, Türkiye, ndogan@ymail.com , 0000-0002-8267-8469

to the studies. There are environmental factors that directly or indirectly affect human health. The main environmental factors are water and food. Water, which is necessary for the continuation of vital events, should not contain harmful chemicals and small disease-causing organisms. In regions with unfavorable infrastructure, water can become a dangerous carrier for health. It is of great importance to analyze the water and to determine the total dissolved substance (TDS) amount in it to learn that it is healthy from the source to the stage when it is consumed. An incentive system is needed to analyze water in a widespread and sustainable way at the point of consumption.

In their research, Muammer Akgün et al. said, "One of the most important problems of 20% of the world's population is the problem of access to clean drinking water. The records of the United Nations show that 1.2 billion people in the world do not have access to clean drinking water and another 2 billion people have to use polluted water. One of the most important problems in developing countries is water-borne diseases caused by polluted surface waters. It is estimated that a lack of clean water affects the health of 1.2 billion people and causes the death of 15 million children annually worldwide" [7]. The total amount of dissolved substances in water resources, briefly TDS, originate from natural resources, urban and agricultural wastes, sewage, and industrial wastewater. Salts used to remove ice from roads in cold weather conditions can cause TDS increase in water resources [8]. In her study, Ayşe Baysal said, "Keeping the intracellular and extracellular water in balance depends on the adequacy of sodium and potassium. To keep the daily sodium and potassium intake in balance, 3 g of salt should be taken for 10 glasses of water" and stated the daily salt consumption values [9]. Since the amount of salt in the consumed water will increase the total amount of dissolved substances (TDS), it will be possible to prevent exceeding the amount of salt that should be taken daily by determining the amount of TDS. Indonesia has set a maximum value of 500 ppm as the TDS standard by the Ministry of Health. The Indonesian National Standard (SNI) has published it as 6.6 - 9.0 for the PH value of clean water. Based on this problem, water quality measurement and filtration tools were made using Arduino, a PH meter, and a TDS sensor [10]. Considering public health, it is necessary to analyze the water coming to the point of consumption and to encourage individuals who will provide this analysis data. Blockchain technology can be suitable for the incentive system. Michael Paul Kramer et al. have defined blockchain as a decentralized database whose data is distributed over a network and a consensus mechanism is used to verify transactions. This data can be transaction logs or executable code and programs. They explained that the data is processed and added to the chain in blocks [11]. Proof-of-stake (PoS) algorithms developed as a solution to the energy consumption problem of blockchain networks usually sell previously created coins/tokens [12]. Filiz Yüksel defined in her study about tokens as "Tokens represent any asset other than crypto money and provide certain rights

to the holder" [13]. Tokenizing assets means digitizing every asset, real or virtual. Since these tokens belong to whom and the transfer transactions are recorded on the distributed ledger, they cannot be changed and are transparent [14]. There has been rapid development in the metaverse and P2E games, which investors and players have been flocking to lately, and as a result, virtual property sales and coin prices are breaking new records [15]. Developed on top of the play-to-win (P2E) model, "move-to-win" (M2E) models have become popular recently. The STEPN project, the pioneer of the M2E model, monitors the movements of its users with purchased NFTs (an immutable token) and rewards them with an in-app token called Green Satoshi Tokens (GST) in exchange for their physical activity. With the popularity of M2E systems, the concept of the Internet of Things has begun to come together with crypto assets. In addition to the useful business models provided by the Internet of Things devices, users can be offered the opportunity to acquire crypto assets by integrating with the M2E (Move to Earn) model. Among these models, the Helium project aims to create an independent wireless network using LoRa technology. It serves around the world with approximately 30 thousand hotspot devices. Users are required to procure a Helium-approved hotspot device to mine in this project. With the EnviDa project, a decentralized solution can be developed to collect sensitive, future-oriented environmental data [16]. Based on the EnviDa ecosystem, environmental data will be collected in an unmanipulated way, free from political and industrial influences, and will provide the data basis for future environmental projects. In the EnviDa project, a cryptocurrency mining device in the proof-of-work (PoW) architecture is carried out with the mining device called DriveMining, and data is collected through the connected sensors. In exchange for the collected data, EnviDa tokens are earned in a proof-of-stake (PoS) network [10]. In the Planet Watch project, each sensor produced detects the concentration of air pollution in an area. It aims to help the community and city authorities analyze where higher pollution levels are found. Temperature, humidity, carbon dioxide, toxic gases, and particle levels are detected by sensors. The sensor data is transferred to the blockchain network and informs the user with its mobile interface. It provides its existing infrastructure by using the incentive mechanisms behind it (PLANETS Token / Token) to create a global sensor network. When the studies are examined, it has been understood the importance of establishing an incentive system to eliminate the risks that water may face from its source to its consumption, and it has been seen that incentive systems can be provided with blockchain technology.

This study aims to determine the total amount of dissolved substances in the water reaching the end consumer. In return for this data obtained by the consumer and provided to the people around him, the token named WTDS, produced in the energy-saving proof-of-stake (POS) Algorand blockchain network was earned. For this purpose, water analysis data is transferred to the server with a TDS Meter and an IoT device, and in return, WTDS token gain is provided with a model suitable

for the M2E model, which is a sub-model of the P2E model. 80% of the 10 billion produced WTDS tokens are planned to be distributed over 25 years. Daily 876,712 WTDS tokens are allocated and transferred to accounts that send data hourly at the end of the day. The user can access the obtained water analysis data and the gains obtained in return for these data via the mobile application interface. In this study, the server application was carried out with SpringBoot and the mobile application with Flutter, and the users who provided TDS data were rewarded with WTDS tokens in the Algorand network and successfully concluded.

MATERIALS AND METHODS

In this study, the water analysis data collection process is provided by the NodeMCU development board and the TDS Meter connected to it. A TDS meter is a device that displays the total dissolved solids (TDS) of a solution, ie the concentration of dissolved solid particles. TDS displays per million (ppm); 1 ppm indicates 1 milligram of dissolved solids per kilogram of water [17]. The TDS Meter has a measuring range of 0-1000 ppm. The components of the device developed in accordance with the concept of the Internet of Things are shown in Figure1.
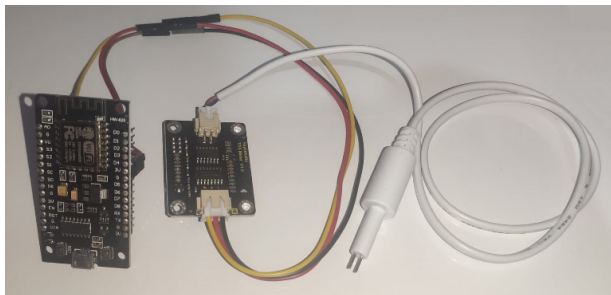
Figure 1. NodeMCU and TDS Meter

The NodeMCU development board broadcasts as a wireless access point. The wireless access point where network scanning is done with the mobile device is shown in Figure 2.
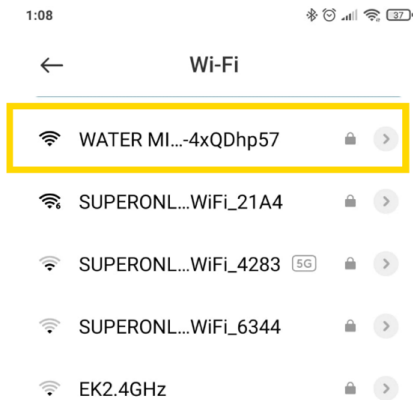
Figure 2. NodeMCU Access Point

By providing a wireless connection to the NodeMCU development board, a wireless network with internet access is defined. The wireless network definition is given in Figure 3.
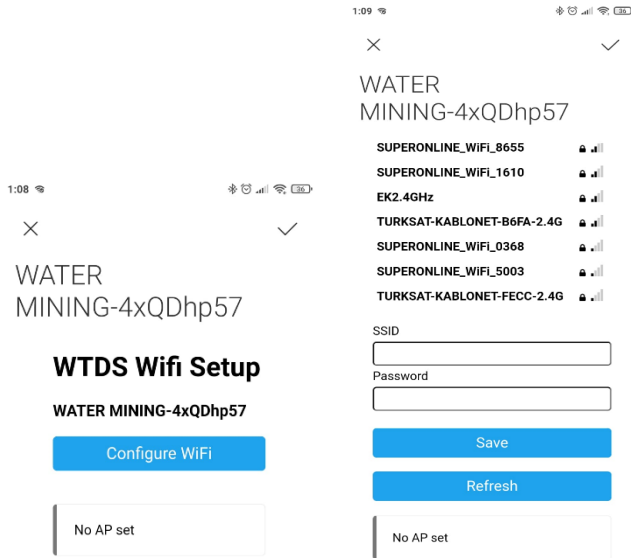
Figure 3. Wireless Network Definition

The device with internet access transfers the collected water analysis data to the server application every hour. The server application is developed in Java with Spring Boot. The server application was developed with the Service-oriented architecture (SOA) approach. Docker was used to quickly compile, test and distribute the developed software. The server application is responsible for retrieving the water analysis data, providing the data needed by the mobile application, and distributing the tokens. The Java package structure of the server application is shown in Figure 4.
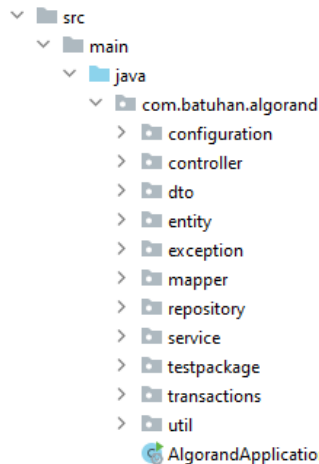
Figure 4. Project Package Structure

Tokenization, which is a part of payment and process development, plays an important role in adapting traditional business models to the blockchain system [18]. Tokens can be produced to maintain the continuity of the platform and increase its quality. At the same time, this unit can represent a value, service, or product. In short, it is possible to say that every product, value, or service turns into a token. The blockchain network used to reward the data collected by the sensors is the Algorand network in the PoS architecture, among the consensus method options. Proof of value (PoS) is an energy-saving compromise protocol alternative to the PoW protocol. Miners in the PoS protocol have to prove ownership of a coin amount (value). People with more assets in this protocol are more likely to be used in verification processes [19]. Blockchain systems working with the PoS algorithm usually sell previously created coins/tokens [12]. When evaluated in terms of high transaction speed and low transaction fee, it has been seen that blockchain applications with high data traffic to be developed, token generation and distribution processes can be carried out efficiently on the Algorand network. Algorand is a decentralized and scalable blockchain network with smart contract generation. It works with an infrastructure called Pure Proof of Stake. Mining is done with the Pure Proof of Stake method, which is an improved version of Proof of Stake and requires less energy than other blockchain networks. In addition to its scalable, secure, and decentralized structure; It aims to provide a solution to the transactional integrity problem experienced in other blockchains. Before branches occur in the block network, transactions are quickly identified and all problems are resolved. With this solution, Algorand contributes to the overall blockchain technology [20]. In this study, 10 billion WTDS tokens created on the Algorand network were produced. 80% of the tokens were developed to be distributed to users for 25 years in exchange for water analysis data. Token distribution is provided to those with IoT devices and mobile application user registration via server application. Mobile application user actions are responsible for device management and display of analytics data. A mobile application has been developed for the crypto asset earnings so that the users are informed of the water analysis data and the incentive mechanism. The mobile application has been implemented in Dart language via the Flutter software development kit, which is suitable for cross-platforms. Flutter is an open-source UI development kit created by Google. It can be used to develop applications for Android, iOS, Windows, Mac, Linux, and the web [21]. The user registration of the mobile application developed in this study and the user operations of the user login screens are given in Figure 5.
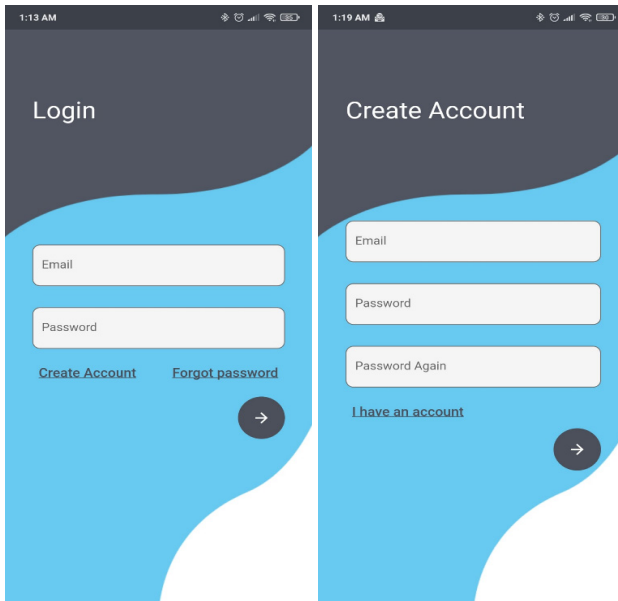
Figure 5. User Operations

Through the user registration screen, each user is provided with the mnemonic information of their wallets that only they can access. The mnemonic information that users need to save to access their personal wallets is given in Figure 6.
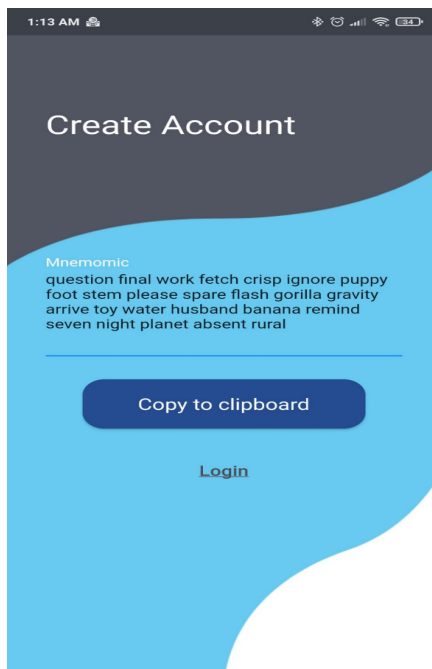

Figure 6. User Mnemonic Information

The user must match with the device that provides water analysis data through the mobile application interface to which the user logs in. The matching process is provided by the device list broadcasting on the same network with the step of adding a new device from the mobile application interface. During user and device pairing, the location of the device is taken with the mobile device and geo-tagged data is collected. User and device pairing is shown in Figure 7.
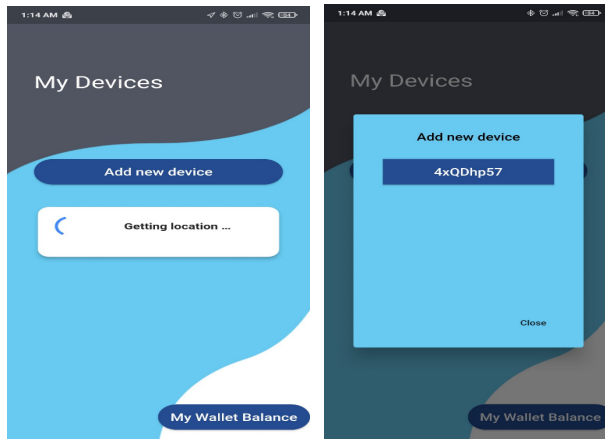
Figure 7. User Device Pairing

Devices belonging to the user are listed on the main screen of the mobile application interface. The latest water analysis data, drinkability status, and estimated coin earnings for each device are shown in Figure 8.
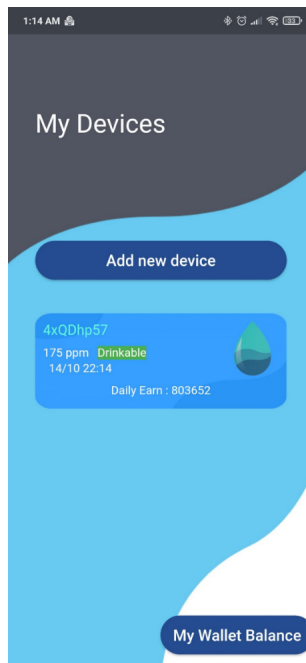
Figure 8. User Device List

In the mobile application interface, it is reported that the water is drinkable in the area where the device providing water analysis data in the range of 50-500 ppm is located. The warning that water not in the range of 50-500 ppm should not be drunk is given in Figure 9.

Figure 9. Undrinkable Water Alert

The 24-hour water analysis data provided by each device via the mobile application interface is shown in Figure 10.
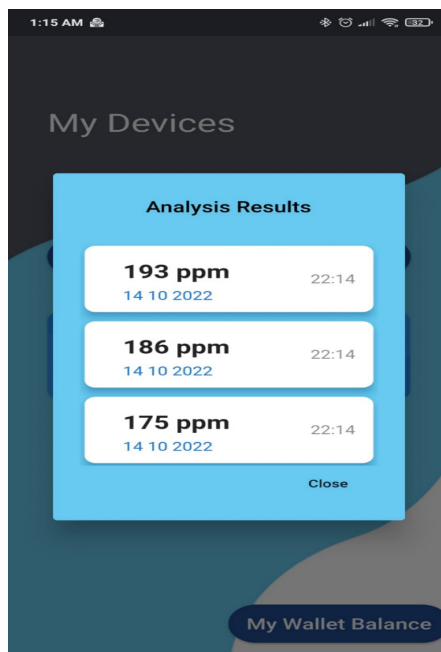
Figure 10. Data of Water Analysis

With the server application developed in this study, token earnings are transferred to the wallets of the users connected to the devices that provide water analysis data at the end of the day. Access to the transferred token earnings is provided with the mnemonic key provided during user registration. The wallet balance that can be accessed with an unrecorded private mnemonic key is shown in Figure 11.
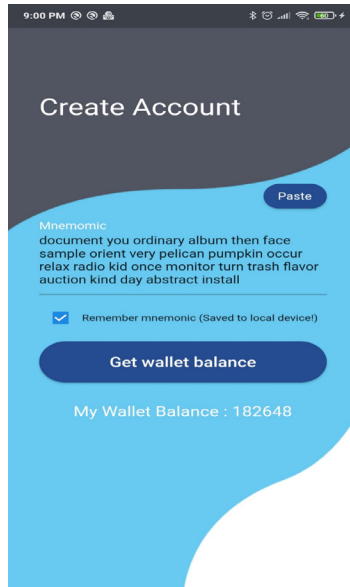
Figure 11. Wallet Balance

The regularly changing water analysis data collection process of the IoT device was provided through the siphon system. The connection of the IoT device to the siphon system is shown in Figure 12.

Figure 12. IoT Device

DISCUSSION AND CONCLUSION

Within the scope of this study, in light of the total amount of solute in water, data that can be linked to the health of people living in the relevant regions will be provided. In terms of timely prevention of sudden water pollution, which causes severe diseases and deaths, its contribution to public health will be high. With the data obtained through this system, local governments will be able to make regional early interventions. It will be able to provide a market opportunity for water treatment systems by detecting water pollution and sharing information. With the IoT devices to be produced and the tokens to be produced for the incentive mechanism, an economic benefit will be provided by bringing and trading the tokens at a price to be determined within the scope of the token economy. With this project, which will be an example of a new crypto asset acquisition model, the meeting model of P2E (Play to Earn) and M2E (Move to Earn) systems with IoT devices, different useful business models will be created. Investment and sustainability of beneficial business models will be ensured by a similar token earning incentive system.

REFERENCES

[1] G. Chen, B. Xu, M. Lu, and N. S. Chen, "Exploring blockchain technology and its potential applications for education," Smart Learning Environments, vol. 5, no. 1, pp. 1–10, 2018.

[2] A. Alammary, S. Alhazmi, M. Almasri, and S. Gillani, "Blockchain-based applications in education: a systematic review," Applied Sciences, vol. 9, no. 12, p. 2400, 2019.

[3] C. Millard, "Blockchain and law: incompatible codes?" Computer Law & Security Report, vol. 34, no. 4, pp. 843–846, 2018.

[4] M. Niranjanamurthy, B. N. Nithya, and S. Jagannatha, "Analysis of Blockchain technology: pros, cons and SWOT," Cluster Computing, vol. 22, no. S6, pp. 14743–14757, 2019.

[5] S. Fosso Wamba, J. R. Kala Kamdjoug, R. Epie Bawack, and J. G. Keogh, "Bitcoin, Blockchain and Fintech: a systematic review and case studies in the supply chain," Production Planning & Control, vol. 31, no. 2-3, pp. 115–142, 2020.

[6] Yuqing Wang, Xueping Han, "Agricultural Productive Service System Based on the Block Chain and Edge Computing", Mathematical Problems in Engineering, vol. 2020, Article ID 6667956, 9 pages, 2020. https://doi.org/10.1155/2020/6667956

[7] "İçilebilir su üretimi için Yenilenebilir Enerji uygulamaları," Termodinamik. [Online]. Available: https://www.yenienerji.com/makale/icilebilir-su-uretimi-icin-yenilenebilir-enerji-uygulamalari. [Accessed: 24-Oct-2022].

[8] WHO. Total dissolved solids in drinking-water. Background document for preparation of WHO Guidelines for drinking-water quality. Geneva, World Health Organization WHO/SDE/WSH/03.04/16, 2003e.

[9] Baysal, A. (2013). Su ve sağlığımız. Beslenme ve Diyet Dergisi, 41(3), 191-193.

[10] Irawan, Yuda & Febriani, Anita & Wahyuni, Refni & Devis, Yesica. (2021). Water Quality Measurement and Filtering Tools Using Arduino Uno, PH Sensor and TDS Meter Sensor. Journal of Robotics and Control (JRC). 2. 10.18196/jrc.25107.

[11] Kramer, Michael & Hanf, Jon. (2021). Is Blockchain the Next General Purpose Technology?. 21. 86-92.

[12] Tikveşli, A. O. (2019). Blok zincir teknolojisi ve% 51 sorunsalı (Master's thesis, Fen Bilimleri Enstitüsü).

[13] Yüksel, F. (2020). KRİPTO VARLIKLAR VE IFRS KAPSAMINDA KRİPTO PARALARIN MUHASEBELEŞTİRİLMESİ . Journal of Accounting and Taxation Studies , 13 (2) , 429-451 . DOI: 10.29067/muvu.653136

[14] Varlıkların Token Haline Getirilmesi [online], halkbank.com.tr, Web address: https://www.halkbank.com.tr/content/halkbank/tr/blog/finans/Varliklarin-Token-Haline-Getirilmesi.html

[15] Vidal-Tomás, D. (2022). The new crypto niche: NFTs, play-to-earn, and metaverse tokens. Finance Research Letters, 102742.

[16] EnviDa çevresel veri takibi ve blokzincir ile mobil madencilik [online], envidatoken.io, Web address: https://envidatoken.io/docs/envida-whitepaper-tr.pdf [Date of Visit: 14 October 2022].

[17] TDS Metre [online], tr.wikipedia.org, Web address: https://tr.wikipedia.org/wiki/TDS_metre [Date of Visit: 14 October 2022].

[18] Ergun, H. & Esenkaya, A. (2022). Blockchain Teknolojisi ile Finansal Piyasalar da Yaşanan Gelişmeler Üzerine bir İnceleme . Karatay İslam İktisadı ve Finans Dergisi , 1 (1) , 77-98 . Retrieved from https://dergipark.org.tr/en/pub/karjief/issue/69419/1101575

[19] Tanrıverdi, M. , Uysal, M. & Üstündağ, M. T. (2019). Blokzinciri Teknolojisi Nedir ? Ne Değildir ?: Alanyazın İncelemesi . Bilişim Teknolojileri Dergisi , 12 (3) , 203-217 . DOI: 10.17671/gazibtd.547122

[20] 6 maddede Algorand (ALGO) Nedir? [online], paribu.com, Web address: https://www.paribu.com/blog/kriptopara/6-maddede-algorand-algo-nedir/ [Date of Visit: 17 September 2022].

[21] Flutter [online], tr.wikipedia.org, Web address: https://tr.wikipedia.org/wiki/Flutter [Date of Visit: 17 September 2022].

T-SER: AN EFFICIENT SPEECH EMOTIONAL RECOGNITION MODEL FOR TURKISH LANGUAGE BASED ON MACHINE LEARNING ALGORITHMS

Emel COLAKOGLU[1], Serhat HIZLISOY[2], Recep Sinan ARSLAN[3]

INTRODUCTION

Emotions have a universal character. However, different meanings can be attributed to emotions in different cultures. In general, the feelings are the same in every society. Emotion is a feeling that indicates that a person is angry, happy, enthusiastic, sad or scared. Emotions are actually states derived from thoughts. People shape their feelings according to the fiction in their minds [1]. Emotions can be observed in the human body as both internal (such as heartbeat, nausea, stomach cramps) and external reactions (such as flushing of the face, opening of the eyes, sweating of the hands, increased respiration). However, the clearest expression of emotion is seen on the face. When we look at a person's face, we can usually understand what mood they are in. Similarly, human voices can be used to understand emotion.

The basis of speech consists of sound [2]. The characteristics of speech are volume, timbre, pitch, duration, tone, emphasis, intersection, melody, and pause. Speech is the most important communication tool between people. For this reason, studies have focused more on the interaction between humans and computers, and audio signals have also been used in this context [3] .

The human voice does not only contain speech data. It also contains a lot of additional information on the audio signal, such as the person's gender, age, dialect, physiology, and mood. Many studies are carried out in order to obtain this information from the audio signal [4-5]. By speaking, people convey not only what they want to say to each other, but also their feelings through sound. Even if we are not face to face with the person we are talking to, we can easily understand what mood he is in from the speech signal [6].

Speech emotion recognition is one of the studies that has started to gain importance in recent years. Speech emotion recognition is used in many areas such as call centers, driving assistance, education, health, entertainment, internet-based learning. Detection of human speech emotions is a rather complex process. Because the analysis of speech signals containing different frequencies and characteristics is quite difficult. Although many different methods have been developed in this area, success rates may vary depending on the quality of the data, language and variety of emotions.

1   Emel Colakoglu, Kayseri University, Graduate Education Institute, Department of Computational Sciences and Engineering, ORCID: 0000-0003-1755-3130
2   Serhat Hizlisoy, Asst. Prof., Kayseri University, Faculty of Engineering, Architecture and Design, Department of Computer Engineering, ORCID: 0000-0001-8440-5539
3   Recep Sinan Arslan, Asst. Prof., Kayseri University, Faculty of Engineering, Architecture and Design, Department of Computer Engineering, ORCID: 0000-0002-3028-0416

One of the most important points in speech emotion recognition is the selection of the data set. The dataset should be up-to-date, diverse, tagged, easy to access, frequently used for comparison, large number of samples, and samples should fully reflect emotions. Initially, a labeled emotion dataset is needed to create an emotion recognition system from speech. One of the factors affecting the success of the study is the labeling quality. There are two approaches to this labeling process. These are called dimensional and categorical. In the categorical method, emotions are grouped with discrete labels. In the emotional approach, the data is shown in the dimensional space with independent axes [1].

There are three main components in the architecture of emotion description systems from classical speech. The first step is to obtain the feature vectors after pre-processing the audio signals [7]. Order to solve a problem with machine learning algorithms, appropriate features must be extracted. But the dataset may not contain ready-made features that can be used. In this case, the features must be removed from the data set. In general, features are collected in two categories in speech emotion recognition processes. These are the Prosodic features Pitch, Energy and Duration and the sound path system (Mel Frequency Kepstrum Coefficient (MFCC), Linear Predictive Cepstrum Coefficients (LPCC)) features. Among the extracted features, the ones most suitable for the problem are selected and given to the machine learning algorithm for the learning of the model. In the learning process, the hyper parameters of these algorithms are optimized with a validation set.

When the previous studies are examined, it is seen that different emotion classes are used in the process of emotion recognition from speech. While most of the studies focus on 4 emotion analysis, there are analyzes up to 7 emotions. In general, high quality datasets prepared for emotion analysis were preferred in previous studies, and therefore, emotion recognition success rates reached high levels. The most used of these datasets are EmoDB, IEMOCAP and RAVDESS. In this study, however, we created our own dataset. The 5-second recording samples we took from Turkish movies and TV series were our disadvantage compared to ready-made datasets due to both sound quality and emotional intelligibility, since there was no shooting in a special environment. However, achieving high recognition rates in this way was a situation that would differentiate the study. In addition, fixed speakers were included in the ready datasets and the training and testing process progressed as speaker dependent. Of course, this is a situation that positively affects performance in previous studies. The second point that differentiates our study is that we have a speaker-independent data set. Very few studies have been conducted independently of the speaker, and success rates have been observed to be low. In this study, a very high success rate was achieved with 85.1%.

Again, although different features were preferred in previous studies, the most used features are MFCC, Energy, pitch, Spectogram, ZCR, LPCC and HNR. While

the singular effects of these features were examined from some studies, hybrid datasets were created from these data from some studies. MFCC attribute was used in our study.

In the classification process, there are studies in which the performances of previous studies were compared using one-by-one classification algorithms. There are also studies in which hybrid models are preferred in order to achieve higher recognition rates. In this study, 11 machine learning algorithms were tested one by one and their results were compared. When we look at the past studies, it has been seen that Naive Bayes, SVM, Decision Trees, Logistic regression, K-nearest neighbor algorithms are mostly used in the process of speech emotion recognition.

RELATED WORKS

Guo et al. (2022) [8], Acoustic features include phase and magnitude information. However, standard speech emotion recognition methods focus only on magnitude information and ignore phase data, thus missing some information. This study aims to extract the phase properties correctly and to use them effectively. The datasets used are EmoDB and IEMOCAP. The features used are MGDCC, DRP, RP and size spectrogram. 2 models have been developed. SCM is the single-channel model and MCMA is the multi-channel model. ELM are extreme learning machines. KELM is a modification of the original ELM and prior studies show that KELM has better performance than ELM. In this study, KELM is used as a classifier. CNN was used to extract phase and magnitude related features from the V1 (magnitude spectrograms), V2 (MGDCC) and V3 (DRP). The highest success rates for both EmoDB and IEMOCAP were achieved with the multi-channel model MCMA. Features included in this success rate are Magnitude, MGDCC, and DRP. For EmoDB (WA 94.02% UA 93.66%) and for IEMOCAP (WA 57.58% and UA 58.07%)

Zhao et al. (2021) [9], IEMOCAP and FAU-AEC were used as datasets. The features are log-Mels, delta and delta-delta. Conv combined with parallel 2D SENet is used as feature extraction block. (PCNSE model). SADRN (Self-Attention Dilated Residual Network) has been used to model long-range dependencies. The CTC layer with the CTC model is used to automatically align emotional labels to what is emotionally salient. Two models were created. PCNSE-SADRN-CTC and PCNSE-SABLSTM. The difference between the two is that instead of the DRN layer in the first, the BLSTM layer is in the second. The best results for the IEMOCAP dataset were obtained with PCNSE-SADRN-CTC. (73.1%). For FAU-AEC, the best results are PCN-SADRN and PCNSE-SADRN-CTC. (41.1%).

Senthilkumar et al. (2021) [10], RAVDESS, IEMOCAP and Emo-DB and are the datasets used in the study. Spectogram is used as feature. Different types of CNN are used for feature extraction. The classifier is a multilayer Bi LSTM. As a result, in both speaker-independent and speaker-dependent situations, the highest success rate in all 3 data sets was achieved with the model developed in this study.

In addition, among the data sets, EmoDB (86% speaker independent) and EmoDB (85.5% speaker dependent) have the highest success rates.

Abdulmohsin et al. (2021) [11], the datasets used are Emo-DB, RAVDESS and SAVEE. The features used in this study are Energy, Harmonic Ratio, ZC, Entropy, MFCC, GTCC, Pitch, Loudness, Fitness, Fourier, deviation of energy, Zero-transition deviation. F-score feature selection was used for feature selection. One hidden layer neural network with 10 nodes is preferred as a classifier. In this study, the most successful classification rates were 86.1% (RAVDESS), 96.3% (Emo-DB) and 91.7% (SAVEE).

Ntalampiras (2021) [12], Log-mel spectrogram and temporal modulation features were used as features in the study. The data set used is Emo-DB. Siamese Neural Network (SNN) was used as a classifier. In this study, the best recognition rate was 82.1% with 9 speakers. In this study, emotional characterization of parts of speech is carried out through analogies, that is, by evaluating the similarities and differences between new and known recordings.

Hizlisoy and Tufekci (2020) [13], they used Turkish emotional data set in their studies. Features were obtained using OpenSMILE, MIRtoolbox and JAudio tools. In feature selection, correlation-based feature selection method was preferred. The WEKA application was used in the classification process. In classification, results from ordinal minimal optimization, logistic regression, decision trees and Bayesian networks were obtained and compared. 10 times cross validation is preferred. The best result was obtained as 95.96% with the combination of MIRtoolbox+Jaudio+OpenSMILE+Bayes. In addition, while the best result was obtained with SMO before feature selection; after the selection, it is provided with Bayesian networks.

Yao et al. (2020) [14], they used IEMOCAP as the dataset. The recognition process is planned for 4 emotions (angry, happy, normal and sad). In this study, a hybrid structure was preferred. In this hybrid structure, DNN, CNN and RNN with different input types are used. With these inputs, the model is trained separately and then the decision-level fusion strategy is used to reach the result. Librosa was used for MS extraction and openSMILE was used for LLD and HSF extraction. The best result was obtained with the recommended fusion method and 57.1% WA and 58.3% UA were achieved.

Langari et al. (2020) [15], in this study, a framework for speech emotion recognition based on an adaptive time-frequency feature extraction using the FRFT is proposed. DFrFT and Cepstral features were used as features. Cuckoo Search and Genetic Algorithm was used for feature selection. 2D CNN is the classifier used in the study. PDREC, EMO-DB and SAVEE were used as datasets. The best results on the basis of dataset (accuracy rates WA, UA) are EMO-DB(97.57%, 97.21%), SAVEE(80%,80%), PDREC (91.46%, 87.31%).

Jia and Zheng (2021) [16], IEMOCAP and custom SER were used as data sets. MFCC is selected as the attribute. The three-layer BiMLSTM model was applied as a classifier. Five cross validation was performed on both datasets. 80% of the dataset is reserved for training and 20% for testing. In both datasets, the best results were obtained with the personalized model + the two-channel general model. (78% (SER) and 70% (IEMOCAP)).

Ancilin and Milton (2021) [17], EMOVO, Berlin, eNTERFACE, Ravdess, Urdu and Savee were used as data sets in this study. The features used are MFCC, MFMC, LFPC and LPCC. In this study, they changed the extraction process of MFCC. They extracted MFMC, which recognizes emotions with better accuracy than conventional spectral features, including MFCC and log frequency power coefficient. SVM was used as a classifier. During the classification process, 10 times cross-validation was performed. The MFMC feature outperformed the other features in emotion recognition in all six datasets. It also recognizes individual sentiments of datasets, with higher recognition rates than LFPC, LPCC and MFCC. The best results were 81.50% for Berlin, 64.31% for Ravdess, 75.63% for Savee, 73.30% for EMOVO, 56.41% for eNTERFACE and 95.25% for Urdu databases. (With MFMC).

Gokalp and Aydin (2021) [3], TESS and RAVDESS are included as datasets in the study. Resnet50 (type CNN), AlexNet (type CNN), Squeezenet (type DNN) and MobileNet (type CNN) were used as artificial neural networks. The features are MFCC and Spectrogram. Classifiers are Decision trees and support vector machine (SVM) and automatic encoders. The best result was obtained with the combination of Alexnet + Auto Encoder on the TESS dataset, with a 98% success rate.

Ozseven (2019) [18], the dataset used in the study is EmoDB. There are 149 features in the study. Of these features, 45 are prosodic and 104 are spectral. MFCC features were obtained with openSMILE and others with Praat. Outlier detection (ADA), forward selection (IDS) and principal component analysis (PCA) were used for feature selection. SVM and k-NN were preferred as classifiers. The best result was obtained with the combination of z-score normalization + MCA classifier + high pass filter + ADA feature selection. (90.3%)

Bhavan et al. (2019) [19], MFCC, Delta and Delta–Delta MFCCs, Spectral Centroids are used as features. Feature selection was also made with Boruta. RAVDESS, IITKGP-SEHSC and EmoDB were used as datasets. SVM is the preferred classifier in the study. The used bagging community consists of 20 SVMs. The training and test data distribution is 90 to 10. 10 times cross-validation was performed on the data. The best result was obtained in the EmoDB dataset with 92.45% (test data).

Huang et al. (2018) [20], the features used in the study are zero-crossing rate, root-mean-square, fundamental frequency, Harmonic-Noise-Ratio and MFCC. MHMC (Multimedia Human Machine Communication emotion dataset) and CHI-

MEI mood disorder dataset were used as dataset. CNN and LSTM took part in the classification process. The best result was obtained with the combination of HSC + EP + LSTM. (75.56%)

Issa et al. (2020) [21], the features used in the studies are Mel-frequency Kepstrum Coefficients (MFCCs), Mel-scaled spectrogram, Chromogram, Spectral contrast feature and Tonnetz representation. RAVDES, EMO-DB and IEMOCAP are the data sets used. The Libora sound library was used for feature extraction. In classification, 1D CNN was preferred. In the study, different analyzes were made and these analyzes were named differently. The analysis, called Model B, had the best results. (96.34% dataset EmoDB).

Aouani and Ayed (2020) [22], 39 MFCC, Zero crossing ratio (ZCR), Teager Energy operator (TEO) and Harmonic noise ratio (HNR) are used as features. Feature selection was made with an auto-encoder (AE). In the classification process, SVM was used. The dataset is the RML sentiment database. The best result was obtained with Simple Auto Encoder as 74.07%. In addition, it was observed in the results that the performance increased after the feature selection process.

Wang et al. (2020) [23], the features used in the study are MFCC and WPC. Analysis was performed on EMODB and EESDB data sets. In order to realize emotion recognition independent of the speaker, the characteristics of an emotion were extracted using wavelet packet analysis. Sequential Floating Forward Search (SFFS) is used for feature selection. The classification process was carried out with SVM. The best result was obtained with the combination of WPC+SFFS+RSVM 79.5% (EMODB).

Pourebrahim et al. (2021) [24], in this study, labeled and unlabeled data were used together. In order to recognize speech emotion on this data, a semi-supervised learning model based on automatic encoders was tried to be created. The datasets used are IEMOCAP, FAUAEC, PESD, EmoDB and RAVDESS. The SSAE method is an automatic encoder-based structure proposed in this study. The proposed method (SSPSE) achieved 61.6% success.

Zhao et al. (2021) [25], in the study, log-Mel spectrogram, delta and delta-delta were used as features. IEMOCAP data set was used. Models were trained for 100 cycles. TCN with 5-layer 1D convolutional modules was used as classifier and LSTM was used for decoder. For comparison, the BLSTM network was preferred. The best results were WA (65.0%) and UA (66.1%) in the test set and WA (68.6%) and UA (69.5%) in the training set. These results were obtained with the self-attention-based TCN model.

Atila and Sengur (2021) [26], MFCC, coleagram and fractal dimension were used as features in their studies. Classification was done with 3D CNN-LSTM. The datasets used are RAVDESS, SAVEE and RML. Python is used for signal-to-image conversion. The deep model was developed in MATLAB. 10 times cross-validation

was done. Success rates for each dataset are RAVDESS (96.18%), SAVEE (87.5%) and RML (93.2%).

METHODOLOGY

In our study, primarily 58 single films and TV series in Turkish were examined for 4 emotions. From here, 409 sound recordings of 5 seconds were taken. In addition, 403 records were created within the 10 second analysis. Camtasia Recorder application was used while taking these recordings. After recording, cutting and editing of the audio files were done with Camtasia Studio. All audio files are arranged to be single channel and sampling frequency is 22050 in all. The file format is wav.

In the second stage, the feature extraction process was applied. Here, MFCC features are extracted from the data. Afterwards, standardization with normal distribution was applied on the data.
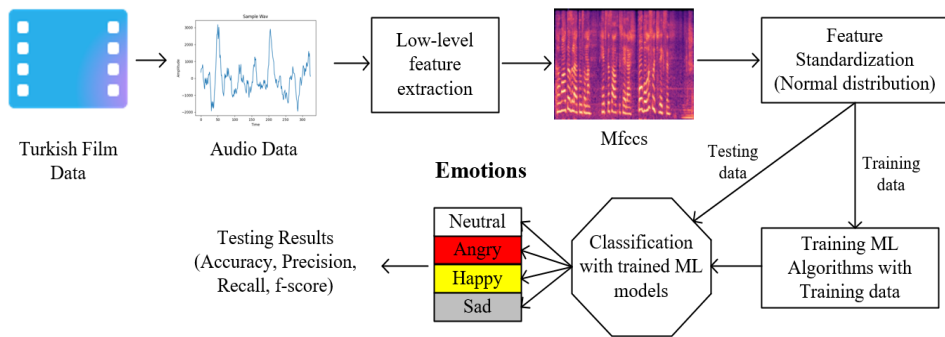
Figure 1. Architectural Structure

The obtained data were randomly separated as training and test data for classification. Different classification algorithms were trained with the training data and subjected to the classification process for the test data. In the classification process, which is the last stage, SVM, Logistic Regression, k-NN, Decision Tree, Gaussian NB, Linear Discriminant Analysis, Ada Boost, Gradient Boosting, Extra Trees, XGB and Random Forest algorithms were tested and the best results were obtained with SVM. As a result, the best performance in emotion recognition is in angry and neutral emotion; the most unsuccessful recognition rate was obtained in the feeling of happiness. Details of all processing steps are given in this section.

Feature Extraction with MFCC

MFCC is a feature extraction technique this is frequently used in speech emotion recognition methods. In MFCC, frequency bands are located logarithmically. MFCC is calculated using Equation 1 [27].

$$Mel(f) = 2595 * \log 10 (1 + \frac{f}{700}) \tag{1}$$

In order to extract the cepstral coefficients, the speech sample is taken as input and the Hamming window is applied to minimize the discontinuity of the signal.

These windows are then used in conjunction with the Discrete Fourier Transform to construct the Mel filter bank. The width of the filters changes according to the Mel frequency curve, and thus the features on the critical band around the center frequency are calculated. Finally, the obtained coefficients are used to calculate the inverse Fourier Transform coefficients [7].

The MFCC feature extraction steps are shown in Figure 2.
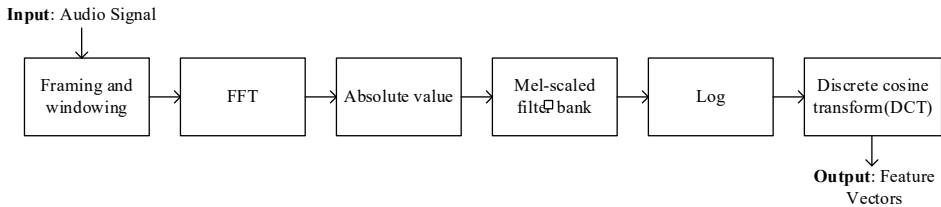
Figure 2. MFCC feature extraction steps [27]

DATASET DETAILS

We prepared the dataset to be used within the scope of the study ourselves. Camtasia Recorder application was used for this process. In addition, all cutting and editing operations on audio files were done with Camtasia Studio. The file format of the audio recordings is wav. All recordings are single channel and sampling frequency is 22050.

At this point, 58 single series of movies were examined for the 5 second data set. In this issue, the sections of the related series are counted as one. With the same setup, 49 TV series were analyzed in a 10-second dataset. As a result of this examination, data in the following breakdowns and numbers were prepared (Table 1).

Table 1. Emotion Distribution for the 5 Sec. Dataset

Angry	Happy	Sad	Neutral
100	96	105	108

In this prepared data, 100 records that were interpreted as the most successful for 4 emotions were selected and the 5 sec. data set consisted of 409 records. 90% was used as a training set and 10% as a test set. For 10 seconds, all 403 data prepared were used. Again, 90% was chosen as the training set and 10% as the test set. When the results obtained in the 5 sec. and 10 sec. datasets were compared, it was determined that there was no significant difference in recognition rate.

While determining the data set, attention was paid to the fact that there was no music in the TV series and movies watched, the sound quality was smooth and at the same time there was only one voice, and there were scenes with a single speaker where the speakers did not interrupt each other. To achieve this, a lot of movie and serial scenes were watched and eliminated.

Each sample used in the data set belongs to only one speaker and a speaker-independent model was created. When we look at the literature, speaker-independent models are seen as a disadvantage for success in such studies. Because in the studies in the literature, generally successful results have not been obtained. This work has achieved success, although the speaker is independent.

STANDARDIZATION WITH NORMAL DISTRIBUTION

Normalization is one of the standard preprocessing techniques. In many studies, min-max or z-score normalization is used in the preprocessing process. In min-max normalization, the features are normalized between [0, 1]. Formula 2 is used for this process [28].

$$v' = \frac{v - min_A}{max_A - min_A} \tag{2}$$

max_A and min_A are the maximum and minimum value of the A attribute. V is the original value of A and v' is the normalized value of A. This way the minimum and maximum values are mapped between [0, 1]. Z-score normalization is calculated using formula 3 [28].

$$z = \frac{v - \mu_A}{\sigma_A} \tag{3}$$

μ_A and σ_A are the mean and standard deviation of the A attribute. V represents the original value of A and z the normalized value of A. After normalization, the mean and standard deviation of all values lie between [0, 1]. The standard normal distribution is actually the z-score normal distribution. In our study, the standard normal distribution was also used.
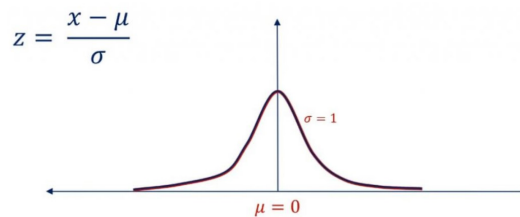
Figure 3. Standard normal distribution [29]

Classification Methods

In the classification process of our study, different algorithms were used and the results were compared.

K-Nearest Neighbors

K-NN classifier is one of the pattern recognition methods that classifies training samples by looking at the closest ones to each other. In this method, it classifies by looking at its k nearest neighbors according to the k value given. In this classification method, a vector is classified based on vectors whose classes are known [30].

K-NN classifier is a statistical learning method proposed by Cover and Hart in 1968. Advantages of the K-NN method; simplicity, efficiency, and high success rates in noisy data. The disadvantages of the k-NN method are that it has low accuracy values in large-sized data, it requires high hardware performance, the uncertainty of the distance type, and high computational cost [31].

In this method, the sample data to be tested is compared with each sample in the training set one by one. In order to determine the class of the tested sample, k pieces of values closest to that sample are selected in the training set. If the number of these selected samples is higher, it is said that the tested sample belongs to this class. The distance between samples is found by the Euclidean (Euclidean) distance. It is stated in the literature that the most appropriate k value is taken by trial and error and generally as 3, 5 or 10 [30].

Logistic Regression

Logistic regression is a method used to classify data. Similar to standard regression analysis, it investigates the relationships between independent and dependent variables. It is a flexible, easy-to-use method that can achieve successful results with few variables [13].

The basic point of logistic regression is to create a regression equation to determine which group the data to be classified belongs to. The subject of logistics modeling dates back to the 1845s. It is observed that the use of this analysis in military matters, education, internal migration movements and meteorology has increased in recent years. Still, the most common use area is in health. Logistic regression has an important place in categorical data analysis. It comes from the logit transformation applied to the dependent variable of the logistic regression analysis. The resulting model (intended model) obtained as a result of logistic regression analysis is a nonlinear function [32].

While the value of the dependent variable is estimated in the regression equations, the probability of realization of one of the values that the dependent variable can take is estimated in the logistic regression. The main feature that distinguishes logistic regression from linear regression is that the outcome variable is binary [33].

Decision Tree

Decision trees are used in both classification processes and regression problems in machine learning algorithms. The basic logic of the decision tree is as that. For a value whose class is unknown, one or more consecutive decision functions are used and the class it belongs to is determined as a result [30].

In general, the decision tree consists of a root node and multiple internal and end nodes. Root and internal nodes are non-terminal nodes and connect to decision stages. Terminal nodes represent final classifications. The root node represents the set of all classes in the process. A set of nodes at a given level in the tree, that is,

the same "distance" from the root, is called a layer. Each node consists of a set of classes to be distinguished, a set of features to be used, and decision rules necessary for classification [34].

Support Vector Machines (SVM)

SVM is a method based on the principle of minimizing structural risk and statistical learning theory. It is used to solve regression and classification problems. SVM solves any classification or regression problem by transforming it into a quadratic programming problem without being stuck with local solutions. The biggest advantage of SVM over other algorithms is that it does not get stuck in local solutions. Also, the generalization ability of SVM is very good. SVM has been used as a classifier in many different studies. Examples of these are face recognition, voice recognition, 3D object recognition, text classification and handwriting recognition [35].

Parameter optimization and kernel function selection have an important place in SVM models. In the applications in the literature, it has been seen that the radial basis kernel function is used with the thought that it generally gives better results [36]. The main purpose of SVM is to obtain the most suitable hyperplane that can separate the classes from each other. That is, it tries to maximize the distance between vectors belonging to different classes.

SVM was introduced in 1995 for the solution of pattern recognition problems. SVMs were developed to figure out the classification problem, but have also been extended to figure out regression problems. The main purpose is to create a dividing line or hyperplane between the data of the two classes. SVM is machine learning algorithms developed to solve the two-class and multi-class classification problem. SVM is basically divided into two according to whether the data set can be separated linearly or not. In case the training dataset is linearly separable, the SVM tries to find the separating hyperplane with the largest boundary [36]. The hyperplane linearly separates data of any size. This hyperplane can be two-dimensional, or more. SVM creates parallel partitions by creating two parallel lines. It separates the area in a single pass to create straight and linear sections. SVM tries to find the widest gap, i.e. the largest margin, between the categories, dividing the two categories as wide as possible. This partitioning is done by the hyperplane [37].

Nonlinear SVMs are algorithms used when the data set cannot be separated by a linear function exactly or with a certain error. In real life problems, linear separation of a data set with a hyperplane is often not possible. Therefore, the process of separating the classes is possible by estimating the separation curve. However, in practice, the curve is very difficult to predict [36]. Kernel functions are used to classify data rather than deriving nonlinear curves. SVM can also be used in nonlinear classification tasks with the implementation of kernel functions [37].

Gaussian Naive Bayes Classifier

Naive Bayes is a probability-based machine learning algorithm based on Bayes' Theorem used in different classification problems. Bayes' Theorem in this definition is a formula that calculates conditional probabilities. Naive Bayes algorithms are generally used in recommendation systems, spam filtering and sentiment analysis. The method is quick and easy. However, the need to be independent of the estimators is the biggest disadvantage [38].

In Gaussian Naive Bayes, it is assumed that the continuous values associated with each feature are distributed according to a Gaussian distribution. When the graph is drawn according to the values obtained, a bell curve that is symmetrical according to the average of the attribute values is obtained. This graphic format is shown in Figure 4.
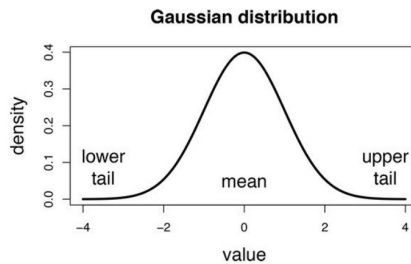
Figure 4. Gaussian distribution [38]

Extra Tree

The Extra Trees classifier is a community tree-based machine learning algorithm. It relies on randomization to reduce computational cost and variance. Extra Trees can be used for classification or regression problems. Extra Trees trains a large number of decision trees. Then, they aggregate the results from decision trees to create predictions. It is similar to Extra Trees Random forest. There is very little difference between them. Random forest uses the bagging method to select different variables from the training data. However, extra trees use the entire dataset to train decision trees. The other difference between the two algorithms is that extra trees randomly select the values that a feature will split and child nodes will generate. However, an algorithm is used for this operation in the Random forest. Also extra trees is faster than random forest [39].

Gradient Boosting Classifier

Boosting is basically a method of transforming weak learners into strong learners. It does this recursively, incrementally. There are multiple boosting algorithms. The main difference between them is how the weak learner defines his/her deficiency. In Gradient Boosting, the first leaf is created first. Then, new trees are created taking into account the prediction errors. This cycle continues until the specified number of trees or more success than expected. First, the target value is averaged. This is the

first leaf to be accepted as a first guess. Then, this value is compared with the target value and it is determined how much estimation error was made. This error value is calculated by subtracting the erroneous value from the observed value [40].

Linear Discriminant Analysis (LDA)

Linear Discriminant Analysis (LDA) is a linear decision boundary classifier created by fitting class conditional densities to the data and applying Bayes rule. This method is also known as the generalization of the linear discriminant. This method converts the dataset to a lower dimensional dataset. In other words, it is the reduction of dimensions in order to reduce model complexity and reduce computational costs. The normal LDA model fits any class with Gaussian density and also assumes that each class uses the same covariance matrix. LDA is closely related to ANOVA (analysis of variance) and regression analysis, which attempt to express a dependent variable as a linear combination of other characteristics or measures [41].

XGBoost

It is a machine learning algorithm based on gradient boosting and decision tree algorithm. It is the optimized performance version of the Gradient Boosting algorithm. It is the best of decision tree based applications. It has high predictive power and speed. It can also prevent over-learning. Higher achievements can be achieved using fewer resources. The working process starts with the first guess. This estimate can be any number as the correct result will be reached by converging with the actions to be taken in the next steps. This number is 0.5 by default. How good this estimation is examined by erroneous estimations of the model. Errors are found by subtracting the estimated value from the observed value. The working logic of XGBoost and Gradient Boosting is the same. There are some differences in detail. They differ in regularization, pruning, working with nulls, and system optimization [42].

ADABoost

AdaBoost is considered the first boosting algorithm. In this model, the education process begins with the training of the weak learner. Incorrectly guessed examples may occur during this tutorial. In the next training, these incorrectly estimated training data are given priority, that is, their weights are increased and retrained. In this way, the output of the weak learner becomes the input to the other learner. At the last point, all results are combined and the final result is obtained. Decision trees with maximum depth trained from the training set are not used. Instead, AdaBoost is used as a weak learner, and decision trees consisting of one node and two leaves with a depth of one are used. Correct ordering is important because it has a cascading structure. The error made in the first decision tree affects the weights of the other decision tree. A sequential computation is used in the AdaBoost algorithm [43].

Random Forest

Random forest algorithm is in supervised classification algorithms. It can be used in both classification and regression problems. Basically, in this algorithm, more than one independently working decision tree comes together and the highest scoring one is selected. In this case, the success rate increases as the number of trees increases. The main difference from decision trees is that the process of finding the root node and splitting the nodes in the random forest is random [44].

The biggest disadvantage of Decision Trees is that they memorize data with excessive learning. To solve this problem, the Random Forest model creates and trains multiple subsets of both the data set and the attributes. If the problem is regression, the average of the predictions of the decision trees, and if the problem is classification, the most successful one is selected. In addition, another important feature of the random forest is that we can give it a certain number of attributes and ask it to choose the most suitable ones among them [45]. We also observed that there are studies using it in the literature.

Evaluation Metrics

To compare the performance of classification metrics, criteria such as accuracy, precision, recall and F1-measure which are frequently used in the literature.

$$\text{Accuracy} = \frac{TN+TP}{SUM} \qquad (4)$$

$$\text{Precision} = \frac{TP}{TP+FP} \qquad (5)$$

$$\text{Recall} = \frac{TP}{TP+FN} \qquad (6)$$

$$\text{F1-measure} = 2 * \frac{Precision * Recall}{Precision + Recall} \qquad (7)$$

RESULTS

In this study, it is aimed to automatically determine emotional states for 4 different classes by using the sounds obtained from Turkish films. For this purpose, tests were performed with different classifiers and the results were evaluated comparatively and shown in detail in Table-3. According to Table 2, it is seen that the best result is obtained with the SVM classifier. In addition, GradientBoosting, ExtraTrees and RandomForest algorithms also showed a recognition performance of over 70%. The dataset was randomly allocated as 90% training and 10% testing. In the confusion matrix, 0 corresponds to angry, 1 happy, 2 neutral and 3 sad. For all emotions, 150 records were used as training sets and 10 records as test sets.

Table 2. Results and Comparison of All Algorithms

Algorithm Type	Accuracy	Precision	Recall	F1-measure
SVM	**82.5%**	**84.2%**	**82.5%**	**82.5%**
Logistic Regression	57.5%	57.3%	57.5%	57.3%
K Nearest Neighbors	45%	54%	45%	44.6%
Decision Tree	35%	34.7%	35%	34.3%
GaussianNB	50%	49.6%	50%	43.3%
Linear Discriminant Analysis	62.5%	63.8%	62.5%	62.8%
AdaBoost	47.5%	53.7%	47.5%	47.6%
Gradient Boosting	70%	70.8%	70%	70.2%
Extra Trees	75%	80.5%	75%	73.9%
XGB	60%	61%	60%	59.4%
Random Forest	77.5%	77.9%	77.5%	77.2%

Gridsearchcv was used to determine the most appropriate parameter values for the SVM algorithm, where the best result is obtained. The best parameter values obtained afterwards are shown in Table 2. In addition, parameter ranges and the default value for the relevant parameter are also included in the same table.

Table 3. Hyper Parameters

Parameter	Used Value	Parameter Ranges	Default Value
probability	True	True, False	False
C	1.0	0.1, 1, 10, 100, 1000	1.0
kernel	rbf	Linear, rbf, poly, sigmoid, precomputed	rbf
gamma	scale	Scale, auto or float	scale
break_ties	False	True, False	False
decision_function_shape	ovr	Ovr, ovo	ovr
class_weight	None	Dict, Balanced, None	None
tol	$1e^{-3}$	Float values	$1e^{-3}$
shrinking	False	True, False	True
degree	6	0, 1, 2, 3, 4, 5, 6	3

When we examine Table 2, there are 3 parameters that we use different from the default value. These are probability, shrinking and degree. Probability means whether to enable probability estimates. The shrinking value determines whether to use the heuristic or not. Degree is the degree of the polynomial kernel function.

In addition, one of the few parameters that is important is gamma. If a low value is selected, the training will sparsely fit the dataset, while a higher value will result in full fit. But in this case, it can cause over-learning.

The Confusion matrix obtained after the classification with SVM is shown in Figure 5 and Figure 6.
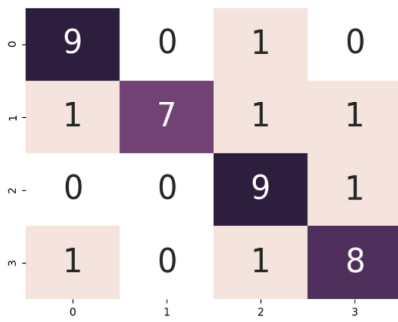
Figure 5. Confusion Matrix
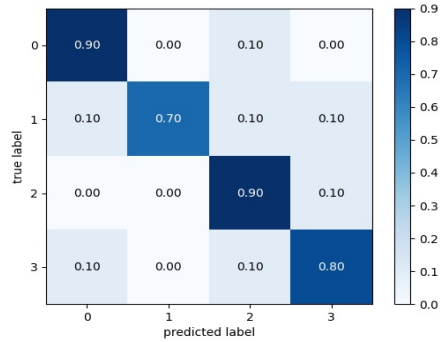
Figure 6. Confusion Matrix (Percentage Value)

According to these two figure, the feeling of angry and neutral is recognized at the rate of 90%, and the emotion of sadness is in the second place with 80%. The lowest recognition rate was 70%, and happy emotion was also obtained. When we look at the general result, the emotion recognition rate is 82.5%. In the training and testing process of other algorithms, the 150-10 distribution in SVM was used. All results are listed in Table 3.

In the third analysis, k-fold cross validation was used. In this method, the data is divided into equal parts according to a determined k value. This ensures that each part is used for both training and testing. Thus, errors and deviations caused by dispersion and fragmentation are minimized [46]. Only in this case, an additional time delay occurs during the training and testing process. Although this is not a problem for small data, it can be costly in terms of computation and time in large data.

The parameters used in K-fold cross validation and the most suitable values are given in Table 4.

Table 4. K-fold Parameters

Parameter	Used Value	Parameter Ranges	Default Value
n_splits	10	Int values	5
n_repeats	3	Int values	10
random_state	123	RandomState instance or None	None

According to Table 4, values different from the default value were given to all parameters and the best results were obtained with the values in the column named used value. n_splits is the card number value. It cannot be less than 2. n_repeats

is the number of times the cross validator should repeat. random_state checks the randomness of all repeated cross validation instances. An integer value is required for reproducible output between multiple function calls.

Table 5. Results and Comparison of All Algorithms (Comparison with K-fold)

Algorithm Type	Accuracy	Accuracy(With K-fold)
SVM	**82.5%**	**85.1%**
Logistic Regression	57.5%	81.3%
K Nearest Neighbors	45%	-
DecisionTree	35%	58.6%
GaussianNB	50%	51.6%
Linear Discriminant Analysis	62.5%	64.5%
AdaBoost	47.5%	53.1%
Gradient Boosting	70%	78.3%
Extra Trees	75%	81.3%
XGB	60%	-
Random Forest	77.5%	80.7%

When we examine the Table 5, an increase in accuracy values was observed in all algorithms after K-fold. Again, the best results were obtained with SVM (85.1%). The highest percentage increase was achieved in Logistic Regression. Figure 7 is the graphical result of Analysis 3.
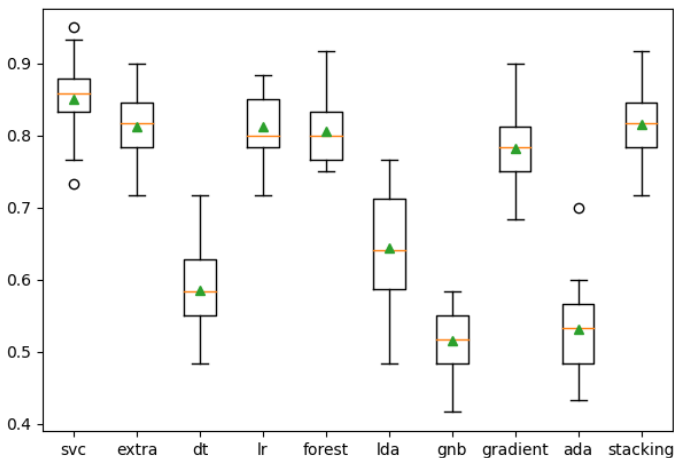
Figure 7. Results of Analysis (K-fold)

In the stacking model, Extra Trees is used in Layer 1 and SVM is used in Layer 2. The accuracy rate obtained in this 2-layer structure is 81.5%. As can be seen here, higher performance has been achieved with single-layer SVM.

DISCUSSION

When the classification algorithms were compared in our study, the best result was obtained with SVM. The accuracy value is 82.5%. In the SVM algorithm, the best recognized emotion was angry and neutral with 90%, while the emotion with the least recognition rate was Happiness (70%). When k-fold cross validation was applied, the recognition rate increased in all algorithms. But again, the best results were obtained with SVM. (85.1%) In a 2-layer model test as a classifier, the accuracy rate remained at 81.5%. (Used Extra Trees in Layer 1 and SVM in Layer 2). In this case, the best result was obtained with SVM classifier (Single Layer) + k-fold cross validation.

Table 6 is the table of result obtained when we compare our model with similar studies conducted recently (in 2020 and 2021).

Table 6. Comparison with Previous Studies

Study	Data Set	Feature Extraction	Classifiers	Accuracy Rate
S.Gokalp and I.Aydin (2021)	TESS RAVDESS	MFCC, spectrogram	Decision Trees, SVM, Automatic Encoder	98.00%
S.Langari et al (2020)	EmoDB SAVEE PDREC	DFrFT, Cepstral features	2D CNN	97.57%
D.Issa et al (2020)	RAVDESS, EMO-DB IEMOCAP	Chromagram, MFCCs, Mel-scaled spectrogram, Tonnetz representation, Spectral contrast feature,	1D CNN	96.34%
H.A.Abdulmohsin et al (2021)	EmoDB RAVDESS SAVEE	Energy, Harmonic Ratio, ZC, Entropy, MFCC,GTCC, Pitch,Loudness, Fitness, Fourier, Haar, deviation of energy,Zero-transition deviation	Neural Network	96.30%
J. Ancilin and A. Milton. (2021)	Urdu eNTERFACE Berlin EMOVO Ravdess Savee	MFCC,MFMC,LFPC, LPCC	SVM	95.25%
L.Guo et al (2022)	EmoDB IEMOCAP	MGDCC, DRP, RP, magnitude spectrogram	KELM	94.02%
N. Senthilkumar et al (2021)	EmoDB RAVDESS	Spectogram	BiLSTM	86.00%
S.Ntalampiras (2021)	EmoDB	Log-mel spectrogram, Temporal modulation	Siamese Neural Network	82.10%
K.Wang et al (2020)	EmoDB	MFCC,WPC,SFFS	SVM	79.50%
N.Jia and C.Zheng (2021)	Custom SER	MFCC	BiMLSTM	78.00%
H.Aouani and Y.B. Ayed (2020)	RML	TEO, 39 MFCC, HNR, ZCR	SVM	74.07%
Z.Zhao et al (2021)	IEMOCAP	3D log-mels spectrogram	A TCN with 1D convolutional modules, LSTM,BLSTM	66.1%

Y.Pourebrahim et al (2021)	IEMOCAP FAUAEC PESD EmoDB RAVDESS	-	Auto Encoder	61.60%
Z.Yao et al (2020)	IEMOCAP	-	DNN,CNN,RNN	58.30%
Our Proposed Method	**Our Dataset**	**MFCC**	**SVM, Logistic Regression, KNN, Decision Tree, GaussianNB, Linear Discriminant Analysis, Ada Boost, Gradient Boosting, Extra Trees, XGB and Random Forest**	**85.10%**

As can be seen in Table 6, a great deal of work has been done in recent years on speech emotional recognition. Different features obtained from different datasets were used in the training of various classifiers and the results were as given. When the results obtained are examined, the recognition rates in the ready-made data sets prepared according to the purpose of the study (recognition of speech emotion) are quite high. (Especially in the EmoDB dataset). In our study, tests were carried out with different classifiers for an original dataset obtained from Turkish films and high classification success was achieved. It is valuable to carry out such a study for the Turkish language and to have a high performance in the end. It is thought that the performance will increase if more data is used in this area.

CONCLUSION

Within the scope of this study, speech emotion recognition was performed using different machine learning algorithms. The most important difference of this study from previous studies is that ready-made datasets, which have high accuracy values and were used frequently in previous studies, are not preferred. We used the data set we prepared ourselves within the scope of the study. Our dataset consists of many speaker-independent examples. Our achievement is better than rates in previous studies in speaker-independent models.

SVM, Logistic Regression, K Nearest Neighbor, Decision Tree, GaussianNB, AdaBoost, XGB Extra Trees, Linear Discriminant Analysis, Gradient Boosting and Random Forest were used as machine learning methods. When the results obtained from all classifiers were compared, the best result was obtained with SVM.

In our future studies, we plan to increase the sample size of the data set (we observed that the performance increased in this way).

REFERENCES

[1] Hizlisoy S., Yildirim S., and Tufekci Z., " Music emotion recognition using convolutional long short-term memory deep neural networks". Engineering Science and Technology, an International Journal, 24(3), 760-767, 2021.

[2] Arslan R.S., Barışçı N., "The Effect of Different Optimization Techniques on End-to-End Turkish Speech Recognition Systems that use Connectionist Temporal Classification", 2018 2nd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), 2018: 1-6.

[3] Gökalp S. and İlhan A., "Comparison of Different Deep Neural Network Models in Emotion Recognition", Journal of Muş Alpaslan University Faculty of Engineering and Architecture, no. 2(1), pp. 35-43, 2021.

[4] Arslan R.S., Barışçı N., Arıcı N., Koçer S., "Detecting and correcting automatic speech recognition errors with a new model", Turkish Journal of Electrical Engineering and Computer Sciences, 29, 2298-2311, 2021.

[5] Arslan R.S., Barışçı N., "Development of Output Correction Methodology for Long Short Term Memory-Based Speech Recognition", Sustainability, 11(15) (2019): 4250-4266. Doi: 10.3390/su11154250

[6] Demircan S., "Various approaches to emotion recognition from speech", Konya, 2020.

[7] Arslan R.S., Barışçı N., "A detailed survey of Turkish automatic speech recognition", Turkish Journal of Electrical Engineering and Computer Sciences, 28(6) (2020): 3253-3269.

[8] Guo L., Wang L., Dang J., Chng E. S. ve Nakagawa S., "Learning affective representations based on magnitude and dynamic relative phase information for speech emotion recognition.", Speech Communication, no. 136, pp. 118-127, 2022.

[9] Yao Z., Wang Z., Liu W., Liu Y. ve Pan J., "Speech emotion recognition using fusion of three multi-task learning-based classifiers: HSF-DNN, MS-CNN and LLDRNN", Speech Communication, no. 120, pp. 11-19, 2020.

[10] Senthilkumar N., Karpakam S., Gayathri Devi M., Balakumaresan R. ve Dhilipkumar P., "Speech emotion recognition based on Bi-directional LSTM architecture and deep belief networks", Materials Today: Proceedings., 2021.

[11] Abdulmohsin H. A., Abdul Wahab H. B. ve Abdul Hossen A. M. J., "A new proposed statistical feature extraction method in speech emotion recognition", Computers and Electrical Engineering, no. 93, 2021.

[12] Ntalampiras S., "Speech emotion recognition via learning analogies", Pattern Recognition Letters, no. 144, pp. 21-26, 2021.

[13] Hızlısoy S. and Tüfekci Z. , "Türkçe Müzikten Duygu Tanıma", Avrupa Bilim ve Teknoloji Dergisi, pp. 6-12, Oct. 2020, doi:10.31590/ejosat.802169

[14] Yao Z., Wang Z., Liu W., Liu Y. ve Pan J., "Speech emotion recognition using fusion of three multi-task learning-based classifiers: HSF-DNN, MS-CNN and LLDRNN", Speech Communication, no. 120, pp. 11-19, 2020.

[15] Langari S., Marvi H. ve Zahedi M., "Efficient speech emotion recognition using modified feature extraction", Informatics in Medicine Unlocked, no. 20, 2020.

[16] Jia N. ve Zheng C., "Two-level discriminative speech emotion recognition model with wave field dynamics: A personalized speech emotion recognition method", Computer Communications, no. 180, pp. 161-170, 2021.

[17] Ancilin J. ve Milton A., "Improved speech emotion recognition with Mel frequency magnitude coefficient", Applied Acoustics, no. 179, 2021.

[18] Özseven T., "The Impact of Pre-processing and Feature Selection Methods for Speech Emotion Recognition", Dicle University Engineering Faculty Journal of Engineering, no. 10(1), pp. 99-112, 2019.

[19] Bhavan A., Chauhan P., Hitkul ve Shah R. R., "Bagged support vector machines for emotion recognition from speech", Knowledge-Based Systems, no. 184, 2019.

[20] Huang K.-Y., Wu C.-H. ve Su M.-H., "Attention-based convolutional neural network and long short-term memory for short-term detection of mood disorders based on elicited speech responses", Pattern Recognition, no. 88, pp. 668-678, 2019.

[21] Issa D., Demirci M. F. ve Yazici A., "Speech emotion recognition with deep convolutional neural networks", Biomedical Signal Processing and Control, no. 59, 2020.

[22] Aouani H. ve Ayed Y. B., "Speech emotion recognition with deep learning", Procedia Computer Science, no. 176, pp. 251- 260, 2020.

[23] Wang K., Su G., Liu L. ve Wang S., "Wavelet packet analysis for speaker-independent emotion recognition", Neurocomputing, no. 398, pp. 257-264, 2020.

[24] Pourebrahim Y., Razzazi F. ve Sameti H., "Semi-supervised parallel shared encoders for speech emotion", Digital Signal Processing, no. 118, 2021.

[25] Zhao Z., Bao Z., Zhang Z., Cummins N., Sun S., Wang H., Tao J. ve Schuller B. W., "Self-attention transfer networks for speech emotion recognition", Virtual Reality & Intelligent Hardware, no. 3, pp. 43-54, 2021.

[26] Atila O. ve Şengür A., "Attention guided 3D CNN-LSTM model for accurate speech based emotion recognition", Applied Acoustics, no. 182, 2021.

[27] Bibin Sam Paul S, Antony Xavier Glittas, Lakshminarayanan Gopalakrishnan, A low latency modular-level deeply integrated MFCC feature extraction architecture for speech recognition, Integration, 76, 2021, 69-75.

[28] Jain S., Shukla S. ve Wadhvani R., "Dynamic selection of normalization techniques using data complexity", Expert Systems With Applications, no. 106, pp. 252-262, 2018.

[29] Valchanov I., "Understanding standard normal distribution," 365 Data Science, 20-Oct-2021. https://365datascience.com/tutorials/statistics-tutorials/standardization/. [Accessed: 03-Nov-2022].

[30] Korkmaz O. E., "Emotion Recognition From Speech Signal", Master's Thesis, Trabzon, 2016.

[31] Anggraeni D., Sanjaya W. S., Munawwaroh M., Nurasyidiek M. Y., and Santika I. P., "Control of robot arm based on speech recognition using Mel-frequency cepstrum coefficients (MFCC) and K-Nearest Neighbors (KNN) method," 2017 International Conference on Advanced Mechatronics, Intelligent Manufacture, and Industrial Automation (ICAMIMIA), 2017.

[32] Çokluk Ö., "Logistic Regression Analysis: Concept and Application", Educational sciences in theory and practice, no. 10.3, pp. 1357-1407, 2010.

[33] Atasoy D., "Examination and an application of logistic regression analysis", Master's Thesis, Sivas, 2001.

[34] Swain P. H. ve Hans H. , "The decision tree classifier: Design and potential", IEEE Transactions on Geoscience Electronics, no. 15.3, pp. 142-147.

[35] Eray O., "The Speech Recognition Application With Support Vector Machines", Master's Thesis, Denizli, 2008.

[36] Ayhan S. ve Erdoğmuş Ş., "Kernel Function Selection for the Solution of Classification Problems via Support Vector Machines", Eskişehir Osmangazi University Journal of IIBF, no. 9(1), pp. 175-198, 2014.

[37] Er M. B., "Emotion Analysis in Turkish Music with Machine Learning", Doctoral Thesis, İstanbul, 2019.

[38] Chauhan N. S., "Naive Bayes Algorithm: Everything You Need to Know", KDnuggets, 2022. https://www.kdnuggets.com/2020/06/naive-bayes-algorithm-everything.html. [Accessed: 03-Nov-2022].

[39] Thankachan K., "What? When? How?: ExtraTrees Classifier", Aug 2022. https://towardsdatascience.com/what-when-how-extratrees-classifier-c939f905851c. [Accessed: 03-Nov-2022].

[40] Muratlar E. R., "Gradient boosted Regresyon Ağaçları," Veri Bilimi Okulu, 24-Jan-2020. https://www.veribilimiokulu.com/gradient-boosted-regresyon-agaclari/. [Accessed: 03-Nov-2022].

[41] Sarker I. H, "Machine Learning: Algorithms, Real-World Applications and Research Directions", SN Computer Science, no. 160, 2021.

[42] Muratlar E. R., "XGBoost Nasıl çalışır? neden iyi performans gösterir?," Veri Bilimi Okulu, 24-Mar-2020. https://www.veribilimiokulu.com/xgboost-nasil-calisir/. [Accessed: 03-Nov-2022].

[43] Güzel K., "Boosting Nedir? Adım AdaBoost Algoritması", Medium. https://kadirguzel.medium.com/boosting-nedir-ad%C4%B1m-ad%C4%B1m-adaboost-algoritmas%C4%B1-439cce20ab9a. [Accessed: 03-Nov-2022].

[44] Akdağlı E. "Makine öğrenmesinde random forest algoritması," Medium, 04-Mar-2021. https://ece-akdagli.medium.com/makine-%C3%B6%C4%9Frenmesinde-random-forest-algoritmas%C4%B1-a79b044bbb31. [Accessed: 03-Nov-2022].

[45] Şimşek H. K., "Makine Öğrenmesi Dersleri 5a: Random Forest (Sınıflandırma)", 24 Mar 2018. Medium.https://medium.com/data-science-tr/makine-%C3%B6%C4%9Frenmesi-dersleri-5-bagging-ve-random-forest-2f803cf21e07. [Accessed: 03-Nov-2022].

[46] Şirin E., "Bir bakışta K-fold cross validation," Veri Bilimi Okulu, 08-Aug-2020. https://www.veribilimiokulu.com/bir-bakista-k-fold-cross-validation/. [Accessed: 03-Nov-2022].

A TECHNICAL EVALUATION OF THE PERFORMANCE ANALYSIS OF BAYBURT SOLAR POWER PLANT USING HELIOSCOPE SOFTWARE AND ACCORDING TO IEC 61724

Mustafa Engin BAŞOĞLU[1], Badel Kocagöz DEMİR[2]

INTRODUCTION

Diversity in electrical energy production has gained importance since the beginning of the 2000s. Until the 2000s, electricity generation was mostly based on fossil-based sources in Turkey and in the world (except for hydroelectric-based energy production), while harmful effects on the environment, increased environmental awareness, limited fossil resources, less exposure to global price changes in energy production, etc. [1]. The use of renewable energy sources has been increasing continuously since the beginning of the 21st century. Solar energy, one of the renewable energy sources, has come to the forefront compared to other renewable energy sources due to its advantages such as providing the consumer with the flexibility to produce the energy they need, being scalable and ease of installation [2].

As of the end of 2021, the total power of solar PV plants installed in the world is 849GW [3]. Because of the incentives given in this field in Turkey, especially since 2010, an installed power of 8658.7MW has been reached as of the end of July 2022 [4]. In terms of solar energy potential, our country is in a favorable position in terms of its geographical location compared to many countries, especially the European continent, and some statistical data in this context are summarized in Table 1.

Table 1. Solar Energy Potential Data for Turkey

Definition	Data
Average annual total sunshine duration	2741.07 hours/year
Average daily total sunshine duration	7.50 hours/day
Average annual total radiation intensity	1527.46 kWh/m²-year
Average daily total radiation intensity	4.18 kWh/m²-day

The inability to achieve the desired efficiency level in solar cell technology causes solar-based power plants to operate at low efficiency and capacity. In this

1   Gümüşhane University, Engineering and Natural Sciences Faculty, Electrical and Electronics Engineering Department, Gümüşhane, TURKEY.
2   Gümüşhane University, Engineering and Natural Sciences Faculty, Electrical and Electronics Engineering Department, Gümüşhane, TURKEY.

respect, monitoring the efficiency, performance and capacity factor of PV power plants has a strategic importance in general. In addition, it should be seen as a need in terms of monitoring and reporting the performance of the plant in order to prevent or reduce situations malfunctions. In the literature, the performance of a solar power plant is measured in many case studies. In these studies, performance ratio, mean array efficiency of the solar panels, capacity factor, specific energy, plant payback time etc. Performance of a solar power plant with 3MWp power in Karnataka region in India was evaluated for 2011 data according to IEC 61724 standards. Annual specific energy was obtained as 1372 kWh/kWp and this value was claimed to be satisfactory compared to other countries. In addition, it was observed that the performance ratio was seriously affected due to inverter errors [5]. The performance of the 2.73kW system consisting of 26 multi-crystalline solar panels in the climatic conditions of Muğla province of Turkey has been examined. Data from 2008 was used to determine the system performance. System performance has been evaluated on a monthly, seasonal and annual basis. Accordingly, annual POA radiation, generated electrical energy and average panel efficiency were determined as 1963.9 kWh/m^2, 182.83 kWh and 9.54%, respectively. The performance ratio of the PV system was calculated around 72% [6]. A study investigating the performance of a 1200W PV system using three different solar panels under equal conditions was carried out in the climate conditions of Izmit, Kocaeli, Turkey [7]. In the study, where monocrystalline, polycrystalline and cadmium-telluride type solar panels were used, energy data between October 2013 and December 2014 were used. According to the results, in the system where cadmium-telluride type solar panel is used, it is seen that the solar panel is less affected by the temperature and accordingly the performance ratio is bigger. In addition, POA radiation was measured for the system established around Kocaeli University and a comparison was made with the radiation data of the General Directorate of Meteorology in Turkey. According to the results obtained, the solar energy potential for Izmit is 23.79% higher than the data of the General Directorate of State Meteorology. In the study, which emphasized the 100GW solar PV capacity in 2020, which was put by the Indian government, the design and analysis of a 50kW rooftop grid connected system were discussed [8]. It is stated that the payback period of the PV systems has decreased considerably with the government support and it is 5.7 years in the study. It is thought that roof type systems will become widespread in the domestic market if the state support continues. According to the technical results and evaluations, the months of May and June are the periods when the ambient temperature is high and the efficiency of the solar panel decreases due to drought and less power is produced. Clear skies and low ambient temperatures in September help achieve maximum power generation from the PV plant. The performance analysis of a 5-kWp rooftop PV power plant in Northern India is made and the effect of temperature is discussed. [9]. The daily specific energy was found to be 3.99 kWh/kWp/day. Average daily PV string efficiency,

inverter efficiency and system efficiency were calculated as 11.34%, 88.38% and 10.02%, respectively. Average daily performance ratio and capacity factor are in the order of 76.97% and 16.39%. The annual energy efficiency of the power plant was recorded as 7175.4 kWh. The results show that the energy loss is maximum in May, when the temperature is highest. According to the overall assessment made within this framework, the performance of the installation of PV systems for any region of India is economically and technically comparable. In the literature, the performance of solar panels has been investigated in laboratory conditions or outdoor conditions in systems around 10kW. Performance reviews on large power systems (in the MW range) are usually of a single location or mounting technology. In addition, data is usually one-year. Therefore, in the study [10], performance evaluation of six large PV plants with different locations was made. Different mounting technologies were evaluated under almost the same meteorological conditions. According to the results of the analysis, the performance ratio parameter is mostly as expected with the location and configuration of the power plants. The performance ratio is higher in summer than in winter. However, the mounting type caused a different trend in the performance ratio. The performance ratio change in PV systems with dual axis tracking is larger than the single axis tracking system. Considering the efficiency of the power plants examined, wind speed is seen as another important parameter to be considered in the design process of a PV power plant. Performance and reliability analysis of the first commercial 10 MWp grid-connected on-channel PV power plant in India was carried out under real field conditions [11]. According to the analysis results, the performance of the canal-top PV plant was found to be lower than the ground-mounted PV plants due to the high humidity around the modules despite the low module temperatures. The increase in solar panel temperature, loss of useful thermal exergy over time and high humidity are the factors responsible for the long-term degradation and loss of performance of the PV plant. PV plant installation on water infrastructure is one of the energy production ways that saves the land as well as the water. The performance of the grid-connected PV system in Kuwait, which includes two different PV technologies with 11.15 MW power, was monitored and analyzed based on real data [12]. A total of 5.5MW thin-film solar panels and a total of 5.6MW polycrystalline solar panels were used in the power plant. Data was tracked for 25 months at the power plant in a hot desert environment, thus gaining the necessary experience to prepare the design of the upcoming GW sized PV projects in Shagaya. The performance parameters of the PV system evaluated in this study are final efficiency, reference efficiency, string efficiency, system losses, string capture losses, cell temperature losses, PV module efficiency, system efficiency, inverter efficiency, performance ratio and capacity factor. For thin film and polycrystalline solar panels, parameters such as performance ratio and capacity factor are very close to each other. Losses are greater because polycrystalline solar panels are not more affected by temperature. Most importantly, it has been claimed

that the same performance behavior can be achieved with a similar operation and maintenance strategy in the Shagaya desert environment, since the two PV systems operate at close capacity factor. However, the system with thin-film technology solar panels requires larger land. Two commercial emulators, HOMER Pro and RETScreen Expert, were used to estimate and evaluate the performance of the 20 MW PV power plant located in Adrar, Algeria [13]. This region is known as a very hot region in terms of climate. In the evaluations, it has been observed that high temperatures adversely affect the behavior of polycrystalline silicon cells and the cell temperatures rise above 60 °C in the hottest months and the electricity generation capacity decreases by 40%. Commercial software shows that thin-film solar cells will perform better in hot climate conditions.

In this study, the parameters affecting the energy production performance of a PV system were defined and some simulation studies were made in HelioScope software and the change of these parameters was analyzed. Specifically, the performance of a grid-connected 7.14MW solar power plant located in Bayburt, Turkey, using 2020 energy production data is discussed. The performance of the power plant is based on criteria defined in IEC 61724 standards such as energy production values, specific energy, performance ratio, mean array efficiency, capacity factor. In addition, the simulation of this power plant was made through the HelioScope software and the simulation results were compared with the real data in 2020. The continuation of the article is planned as follows: In the second part, Turkey's solar energy potential and current PV power capacity are presented. In the third chapter, performance evaluation indexes are defined according to IEC 61724 standards. In the fourth chapter, the energy production data and performance of the Bayburt solar power plant for 2020 are examined and discussed. With the analysis performed in HelioScope software, the simulation results were compared with the real data. In summary, a study is presented in which analyzes are carried out on a real application in which the solar energy potential is evaluated for the province of Bayburt.

SOLAR ENERGY STATUS IN TURKEY

Energy, in general terms, is the ability to do work. Energy, which has an important place in our age, shows the welfare level of the society and is accepted as the pioneer of economic and technological progress. The development of the society is in parallel with the existence and amount of energy resources. On the other hand, the demand for energy, which is the basic need of humanity, is increasing day-by-day [14]. The reason for this is the increase in human population, urbanization, industrialization and the growth of trade volume. This increasing demand leads to the depletion of existing energy resources and a tendency towards alternative energy resources. Energy is divided into renewable and non-renewable energy. Non-renewable energy is non-sustainable, exhaustible, and polluting the nature when used. Oil, natural gas and coal are examples of non-renewable energy sources that release enormous

energy because of their burning [15]. Renewable energy is the energy that does not harm the nature and that the resources can renew themselves continuously. Examples of renewable energy sources are solar, wind, water, biomass, hydrogen, geothermal, wave, tidal energy. Renewable and non-renewable energy sources are the sources that we use directly in nature in energy production, also called primary energy. In Fig. 1, primary energy consumption based on energy resources in the world is shown [16].
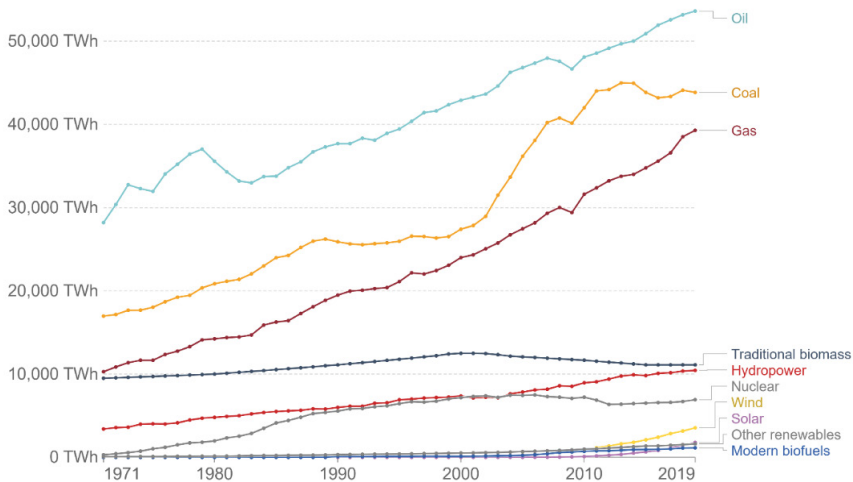
Figure 1. Energy consumption by primary energy sources in the world

According to the Fig. 1, the distribution of primary energy consumption by resources by years, it is understood that the ratio of fossil-based resources such as oil, coal and natural gas is quite dominant. According to 2019 data, the shares of oil, coal and natural gas in primary energy consumption are 30.93%, 25.3% and 22.67%, respectively, and 78.9% in total. Although the use of renewable energy-based resources such as wind and solar has increased in recent years, with the incentive mechanisms and falling costs of the states, it has a share of 3.07% in total consumption. The share of solar energy in total consumption individually has a very low share of 1.03%.

Electric energy consumption in Turkey increased by 7.7% compared to the previous year and reached 329.6 billion kWh in 2021, while electricity generation increased by 8.1% compared to the previous year and reached 331.5 billion kWh [17]. For the period 2020-2040, according to the results of the Turkish Electricity Energy Demand Projection Report study, electricity consumption is expected to reach 370 TWh in 2025 and 591 TWh in 2040, according to the baseline scenario.

The share of renewable energy in electricity generation in our country is increasing every year. Particularly, government of the Turkey policies supporting renewable energy, unlicensed electricity production incentives and satisfactory

incentives per kWh and bonus payments for domestic equipment were effective in the increase in this share. In this context, the share of solar and wind-based electricity generation has increased considerably since 2015. For example, in 2021, 31.4% of electricity generation will come from coal, 32.7% from natural gas, 16.8% from hydraulic energy, 9.4% from wind, 4% from solar, 3.2% from 6% was obtained from geothermal energy and 2.4% from other sources in Turkey. The change in the power of solar PV-based power plants by years is shown in Fig. 2. According to the latest data for June 2022, the total installed power in Turkey is 101.518GW and the installed solar PV power is 8658.7 MW. Distribution of total installed power by resources; 31.1% hydraulic energy, 24.9% natural gas, 20.8% coal, 10.8% wind, 8.4% solar, 1.7% geothermal and 2.4% is in the form of other sources [18]. The share of solar PV's electricity generation in total electricity generation between 2015 and 2020 is given in Fig. 3.
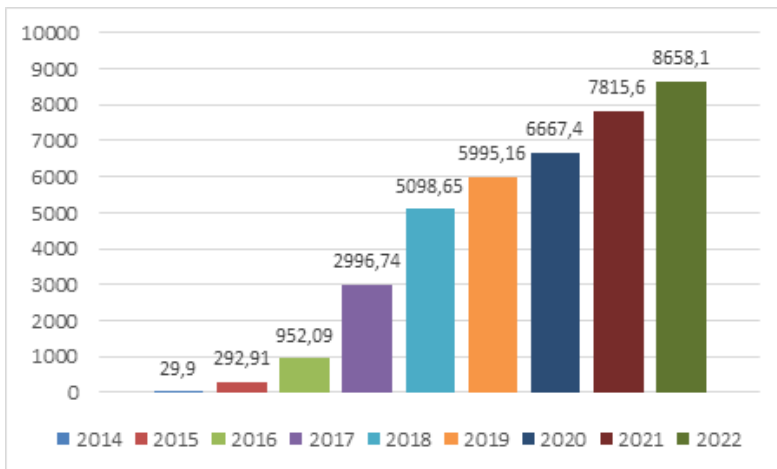
Figure 2. Change of total installed capacity of solar energy plant by year
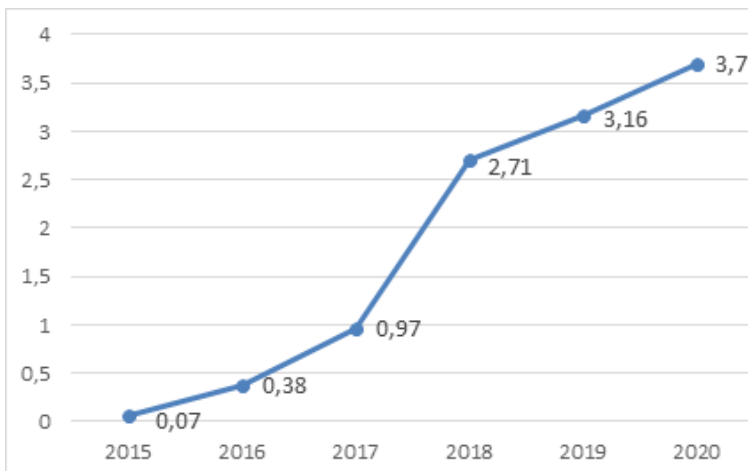
Figure 3. Energy consumption by primary energy sources in the world

PERFORMANCE EVALUATIONS PARAMETERS IN PV POWER PLANTS

The power conversion efficiency of a PV module is very low compared to other electrical equipment in the solar power plant. Performance evaluation of solar panels gains importance as their usable power depends on environmental conditions such as solar radiation, temperature and other weather conditions. To calculate the energy performance of solar panels, IEC 61724, NREL and SMA etc. Some indicators suggested by [26] can be used. In this context, the daily-generated energy in DC or AC form, solar radiation, the nominal power of the solar power plant, the total surface area of the solar panels, the inverter efficiency and the number of days in a month are used in the calculation of performance evaluation metrics. While evaluating the energy production performance in solar power plants, DC energy by the PV module or AC energy transferred to the electricity grid can be taken into account. If parameters in DC form are used in this evaluation, the energy performance of the PV modules is determined. On the other hand, if AC energy is used, the overall energy performance of the plant can be calculated. After these explanations, firstly the monthly produced energy is calculated as in Eq. (1) [6].

$$E_M = \int_0^{t_S} P_A dt \qquad (1)$$

where E_M is the monthly generated energy in AC form, t_s is the sunshine duration corresponding to the inverter running time in a given month and P_A is the output power of the inverter. After calculating the monthly energy, the plane of array solar radiation is calculated. This parameter can also be defined as an energy value per square meter as given below [19].

$$E_{AT}(kWh/m^2) = \int_0^{t_S} Q_A dt \qquad (2)$$

where E_{AT} is the monthly average total PV energy per square meter and Q_A is the instantaneous plane of array radiation in W/m^2. Y_A is the parameter that describes how long the energy produced by the solar panel corresponds to working under standard test conditions continuously. This time can be calculated as given in Eq. (3) [20].

$$Y_A = \frac{E_M}{P_{STC}} \qquad (3)$$

where P_{STC} is the rated or maximum power under standard test conditions. Y_R is the reference time in hours and is calculated as given in Eq. (4) [20].

$$Y_R = \frac{E_{AT}}{Q_{STC}} \qquad (4)$$

where Q_{STC} is the solar irradiance value under standard test conditions. The performance ratio is an important performance evaluation metric for a solar plant

and means measuring the quality of a PV plant regardless of climate or location. As formulated in Eq. (2), the correlation between the real and theoretical energy outputs of the PV plant is defined in Eq. (5) [21].

$$PR(\%) = \frac{Y_A}{Y_R} x100 = \frac{E_M}{\frac{P_{STC}}{Q_{STC}} \int Q_A dt} \qquad (5)$$

Mean array efficiency is another important parameter that determines the power conversion efficiency for one square meter of solar panel surface area. It depends on the size of the PV modules, the energy and solar radiation as given in Eq. (2). This relation is given in Eq. (6), where A is the surface area of the PV arrays and η_A is the mean array efficiency in this equation [19]. In this calculation, the energy must be in DC form.

$$\eta_A(\%) = \frac{E_M}{E_{AT} A} x100 \qquad (6)$$

The capacity factor is the ratio of the total produced energy to the energy that can be theoretical. It can be calculated monthly as given in Eq. (7) [22].

$$C_{F,m}(\%) = \frac{E_M}{24 P_{STC} D} x100 \qquad (7)$$

where D is the day number in a certain month.

BAYBURT SOLAR ENERGY PLANT
Solar Energy Potential in Bayburt Province

The average sunshine duration and average solar radiation intensity values of a region are the most important parameters affecting the efficient operation of the solar power plants to be established. Solar energy-based power plants are systems that produce electricity with variable characteristics depending on the meteorological conditions and the production capacity can be estimated depending on the solar radiation values of their location. Within the scope of this study, according to the data obtained from the General Directorate of Meteorology of Turkey, it can be evaluated that the solar energy potential is high throughout the province of Bayburt. Numerically, the annual total solar radiation value in Bayburt varies between 1500-1750 kWh/m². Radiation values reaching 1700 kWh/m² can be obtained for Demirözü and the central district. Fig. 4 shows the solar energy potential map of the General Directorate of Meteorology for Bayburt [17].
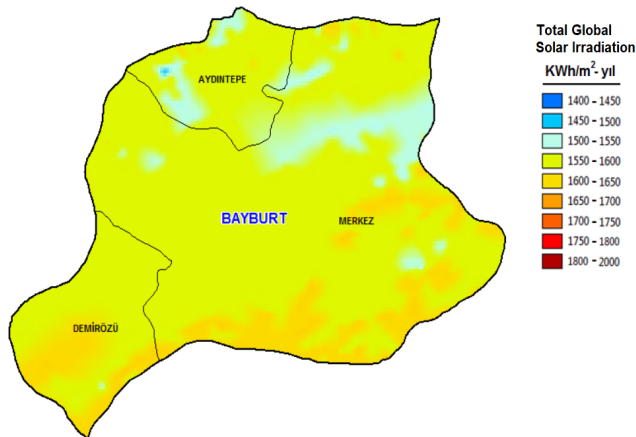
Figure 4. Solar energy potential map of Bayburt, Turkey

Although the estimation of the energy that can be produced for a solar power plant depends on many parameters, it should be considered as a constantly changing situation. These parameters can be summarized as the sum of global solar radiation, average sunshine intensity, ambient temperature and humidity, location, mounting type. According to the data obtained from the Bayburt Provincial Directorate of Meteorology for Bayburt, the total global solar radiation of Bayburt was determined as 224.9kWh/m² in August. In December, the total global solar radiation was measured as 68.9kWh/m². According to the solar energy potential atlas prepared by the Ministry of Energy and Natural Resources, the total global solar radiation variation by months for the city of Bayburt is presented in Fig. 5.
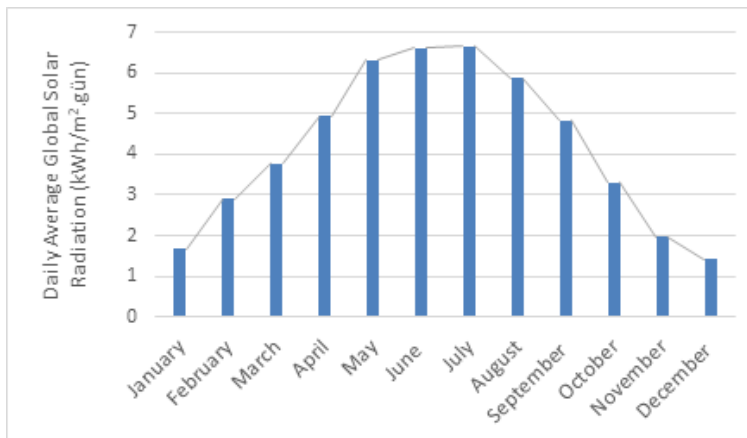
Figure 5. Daily average global solar radiation change by months for Bayburt, Turkey

According to the definition of the World Meteorological Organization, the duration of sunshine is the period of time when the solar radiation directly reaching the unit area exceeds 120W/m² in a certain period. This period can be calculated daily, monthly or yearly for any location. Longer sunshine duration is advantageous

in that solar power plants produce more energy. Longer sunshine duration in summer months enables the power plant to produce more energy. Even if the sunshine duration decreases during the winter months, since Turkey is a middle belt country, the sunshine durations are at reasonable levels for the power plant to operate. According to the measurement results of the Provincial Directorate of Meteorology at the Bayburt station, the total sunshine duration was calculated as 340.4 hours for the month of July at the most. The lowest sunshine duration for Bayburt province was in December with 152.1 hours. In Fig. 6, the change in monthly total sunshine duration in Bayburt province is given.
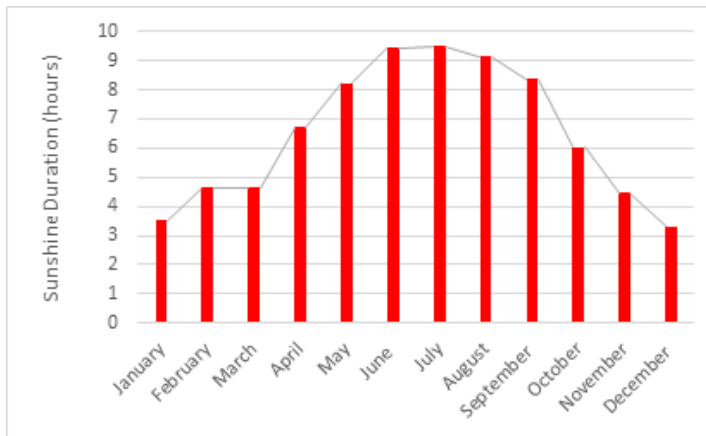
Figure 6. Average sunshine duration by months for Bayburt, Turkey

Bayburt Solar Power Plant

The solar power plant, which is operated by a private company in the Danişment Village of Bayburt province, is connected to the grid and has an installed power of 7.14 MW. There are six hermetic type transformers with a total power of 1250 kVA in the power plant, which is built on a large land. 270W polycrystalline solar panels were used in the power plant. The general view of the solar power plant is shown in Fig. 7. In Bayburt solar power plant, 22 solar panels from a solar array are connected in series. P270 model solar panels placed on the ground are adjusted to face south. The azimuth angle adjusted according to the angle of incidence of the sun's rays is 180°. The inverter used in the power plant is a 60 kW product of SunGrow. These inverters transfer the energy produced by the solar arrays to the alternating current network and are smart inverters that can display some electrical information instantly. In this plant, 16 inverters are connected to each transformer and there are 96 inverters in the plant. The total number of solar panels used in the power plant is 27456.

Figure 7. Image of the Bayburt Solar Power Plant

Annual Performance of Bayburt Solar Power Plant With Real On-Site Data

The energy transferred to the grid by the 7.14 MW Bayburt Power Plant for 2020 has been measured as 12020.75 MWh in total. The monthly energy production change of the plant is given in Fig. 8. The power generation of the power plant reached peak levels in July and August. Energy production and monthly average solar radiation data are listed in Table 2. Solar radiation data were obtained from Bayburt Meteorology Directorate and it was observed that these data contain differences from Bayburt solar energy potential atlas. These differences are also given in Table 2. According to the differences obtained, there are serious differences between the solar energy potential atlas of Bayburt and the year 2020. Since it contains the average of the measurement results taken for a certain period on the map, it may be considered normal for serious differences to occur in some months. In summary, the solar radiation values given for 2020 are 12.23% bigger than the solar energy potential atlas.
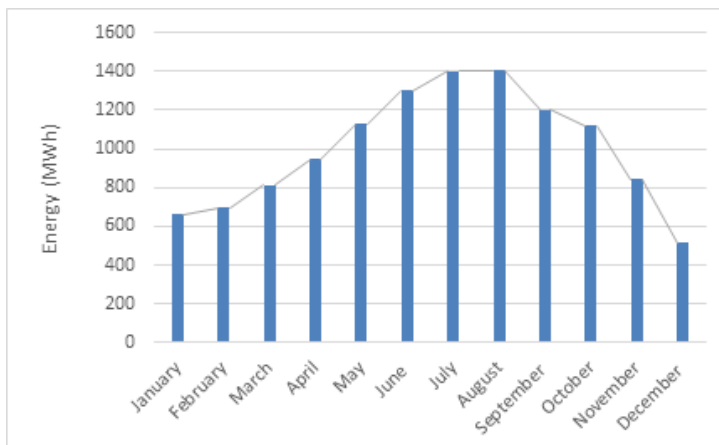
Figure 8. Energy production change of the power plant for 2020 by months

Table 2. Generated energy and average solar radiation for Bayburt

Months (2020)	Generated Energy, 2020 (MWh)	Average solar radiation, 2020 Data Obtained Bayburt Meteorological Station (kWh/m²)	Average solar radiation (Solar Energy Potential Map) (kWh/m²)	Difference (%)
January	661.520	81.5	52.08	36.1
February	693.990	88.8	81.2	8.55
March	812.770	118.5	116.87	1.37
April	944.910	153.1	148.5	3
May	1127.650	166.5	195.61	-17.48
June	1301.170	186.1	198.3	-6.55
July	1398.510	214.5	206.15	3.89
August	1404.330	224.9	182.28	18.95
September	1202.960	174.9	144.9	17.15
October	1117.970	125	102.61	17.91
November	842.070	81.4	59.1	27.39
December	512.900	68.9	43.71	36.56

When the monthly performance ratio values of the power plant are examined, it is seen that the system operates with a high performance ratio in the winter months. On the other hand, lower performance ratio values were calculated in the summer months. According to 2020 data, the performance rate varies between 83-94%. Although the energy produced in summer is high, it is seen that the performance rate in summer months is lower than in winter months. The reason for this is that the temperature-dependent coefficient of the polycrystalline solar panel is high and the losses increase accordingly. Operating with the most efficient performance with 94% in December, the power plant operates with the lowest performance rate with 83% in August. Since the installed power of the power plant calculated under standard test conditions is constant according to the months, the monthly energy value transferred to the grid and the monthly average solar radiation values determined the performance ratio. The variation of the performance ratio calculated according to the months for the Bayburt solar power plant is shown in Fig. 9.
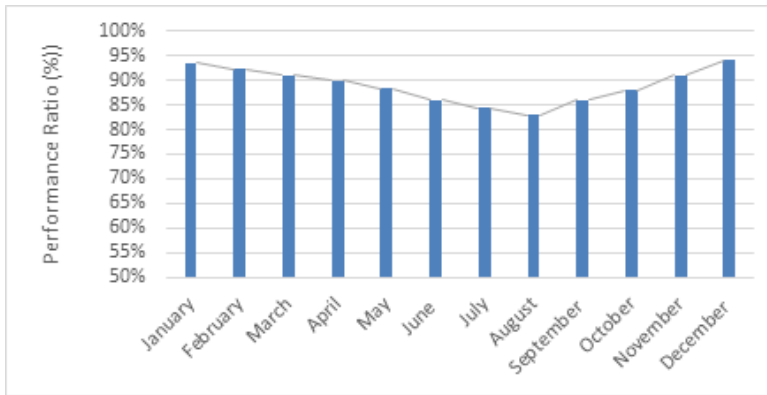
Figure 9. Performance ratio change on a monthly basis for 2020

Since the monthly capacity factor and average solar radiation for Bayburt solar power plant is high, large values are calculated in the summer months. The capacity factor was calculated at the highest value of 25.3% in July and the lowest at 9.3% in December. The capacity factor reached maximum levels in the summer months when the level of solar radiation and sunshine duration are high. However, even under the best conditions, it could not go above the 25% level. The change in capacity factor by months for the Bayburt solar power plant is shown in Fig. 10. Solar panels do not always work at efficiency for standard test conditions. Average operating efficiencies of solar panels in Bayburt solar power plant are given in Fig. 11 on a monthly basis.
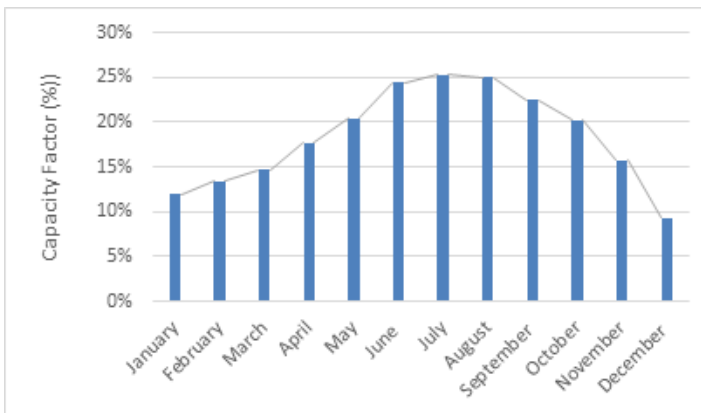
Figure 10. Capacity factor change on a monthly basis for 2020
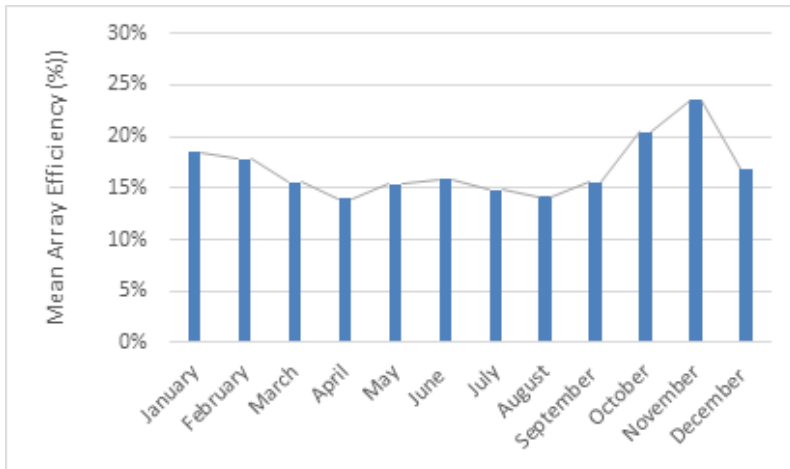
Figure 11. Mean array efficiency change on a monthly basis for 2020

Table 3. Performance metrics of Bayburt solar power plant for the 2020 data

Months (2020)	Performance Ratio (%PR)	Capacity Factor (%CF)	Mean Array Efficiency (%η)
January	93.5	11.99	18.5
February	92.2	13.4	17.79
March	91	14.7	15.61
April	89.9	17.7	14
May	88.3	20.4	15.4
June	86	24.4	15.9
July	84.4	25.3	14.84
August	83	25	14.2
September	86.2	22.5	15.6
October	88	20.2	20.4
November	92	15.77	23.5
December	94	9.3	16.9

Annual Performance of Bayburt Solar Power Plant HelioScope Software

In HelioScope software, a design very similar to the technical features of the Bayburt solar power plant was made and a comparison was made with the power generation data of the plant for 2020. In addition, the performance ratio, efficiency, losses and causes of the losses were evaluated by the analysis.

The solar panels used in the Bayburt power plant are Solar Fabrik brand, with 270W power and polycrystalline structure. In the HelioScope program, solar panels with the closest characteristics available in the solar panel catalog were preferred. It has been ensured that the number of panels placed according to the installed power of the power plant is the same in the design. It is stated in the power plant data that the inverter in the solar power plant is SunGrow brand and has a power of 60 kW.

After the selection of the basic elements that make up the design in the program, the solar panels were placed in the area. The solar panels installed on the sloping land are arranged at intervals so that they do not cast too much shadow on each other. The design was made by leaving the transformer places blank. The image of the solar panels in the power plant is given in Fig. 12.

Figure 12. Design image of Bayburt Solar Power Plant in HelioScope

The total amount of energy produced in Bayburt solar power plant for 2020 is 12020.75 MWh. The annual amount of energy produced by the same power plant designed in the HelioScope program is 11034.620 MWh. It has been observed that there is an 8% difference between the value actually produced by the power plant and the energy production value predicted because of the design. When the monthly energy productions are examined, the plant realized the highest production in August and the lowest production in December, both in reality and in design. The radiation value obtained by the General Directorate of Meteorology for the province of Bayburt is 224.9 kWh/m² for August, and this value has been calculated as 208.2 kWh/m² for the same month in the program. The reason for this is that the HelioScope program calculates the solar radiation by taking the most unfavorable situations as a reference while designing. In Table 4, the amount of solar radiation reaching the location of the designed power plant by months and the estimated energy production values according to the design are listed.

In the analyzes made in HelioScope, the operating temperature of the solar panels is 17.5 °C on average, while the average temperature of the environment where the power plant is located is 8.8 °C. The measured plane of array irradiation in the solar array plane decreased because of shadows falling on the panel. Since the solar panels absorb some of the sunlight they absorb and reflect some of it back, the radiation value decreased a little more after reflection. Pollution and dusting of solar panels have also caused a decrease in the radiation value since they prevent full use of sunlight.

Table 4. Monthly solar radiation and energy production values for Bayburt solar power plant

Months (2020)	Plane of Array Irradiation (kWh/m²)	Generated Energy (kWh)
January	92.9	656.34
February	100 1441	698.6
March	144	961.11
April	156.9	1020.57
May	170	1076.85
June	182.5	1116.44
July	199.8	1212.48
August	208.2	1249.09
September	171	1053.77
October	133.5	862.09
November	88.3	588.37
December	77.4	538.87

The differences between the amount of energy produced by the power plant and the solar radiation values in that region and the annual energy amount to be produced by the designed power plant and the monthly average solar radiation values were evaluated. It is known that these differences are caused by some system losses. One of the system losses is incompatibility between solar panels. Since solar panels are connected in series and form arrays, when any panel is damaged, broken, cracked or shaded, it also affects other series-connected panels. The solar panel that does not operate at the appropriate current due to the shaded panel between the panels connected in series causes incompatibility with other solar panels. Another system loss is related to the operating temperature of the panels. Since the panels will emit heat to the environment while operating, the average temperature around the plant should be low. The irradiance value is another factor affecting the system. The high solar radiation value ensures efficient operation of the power plant. Thus, a high performance ratio value is formed in the plant. Another loss parameter affecting the system is the amount of pollution of the panel. Dusting and pollution of the surface of solar panels prevents the panel from absorbing sunlight. The pollution of the panel and the low utilization of solar radiation are directly proportional. The low amount of light absorbed reduces the performance rate of the plant and reduces the amount of energy produced.

Reflection loss occurs when a panel reflects some of it back without absorbing all the light. When the reflected light is excessive, the performance of the system decreases. Another of the system losses is ghosting. Tree and building shade falling on the solar panel will affect the amount of light absorbed by the solar panels, thus reducing the efficiency of the system. Therefore, the solar panels in the power plant

should be placed in such a way that they are not exposed to shading as much as possible. Inverter losses reduce the performance ratio of the system as they affect the AC system output. Wiring losses are among other system losses. The length, cross-section and type of the cable affect the operation of the switchboard. The losses of the designed power plant and the distribution of losses are shown in Fig. 13.
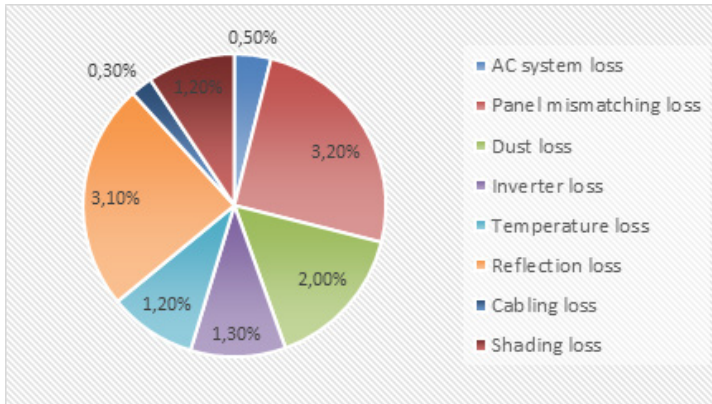
Figure 13. Power loss share with analyzes made in HelioScope software

CONCLUSIONS

In this study, technical information of the equipment such as solar panel and inverter used in the power plant was obtained by examining the field of the Solar Power Plant established in the village of Danişment in Bayburt province. In addition, monthly energy production data for 2020 has been reached at this power plant. From the monthly produced energy values of the Bayburt solar power plant, the amount of energy transferred to the grid annually was calculated and this value was found to be 12.1 GWh in total for 2020.

Performance ratio, capacity factor and mean array efficiency were calculated from monthly average solar radiation values obtained from monthly energy and meteorology for Bayburt solar power plant. The months with the highest performance ratio value were in winter, and the lowest in summer. The plant operating with the highest efficiency in December with a performance ratio of 94%. Due to the summer temperatures and the panel operating temperature, it worked with the lowest efficiency in August with 83%. The annual average performance rate of 89.04% is calculated; Bayburt SPP shows that it works efficiently. The capacity factor calculated at the highest values in the summer season. It reached the maximum value in July with 25.3%. Capacity factor decreasing in winter; It was calculated at the lowest value in December with 9.3%. Average panel efficiency varied by month depending on system losses and temperature. The main reason why it varies according to months is; the angle of incidence of the sun's rays, the monthly average amount of radiation, and the monthly sunshine duration affect the panel efficiency.

Recently, electricity generation from solar energy has become widespread in order to reduce foreign dependency in energy and contribute to the country's economy. Solar energy systems, which provide the opportunity to be installed from small powers to very large powers, offer wide possibilities from building roof-facade applications to land and field installation. The necessary conditions for the efficient operation of the installed photovoltaic systems and the causes of system losses were examined in detail and it was aimed to lead the next studies.

REFERENCES

[1] Sözen A, Arcaklıoğlu E, Özalp M. Kanit EG. Solar-energy potential in Turkey. Applied Energy, 2005;80:367-81.

[2] Başoğlu ME, Comprehensive review on distributed maximum power point tracking: Submodule level and module level MPPT strategies. Solar Energy, 2022;241:85-108.

[3] IRENA (2021), Off-grid Renewable Energy Statistics 2021, International Renewable Energy Agency, Abu Dhabi.

[4] TEİAŞ, Kurulu Güç Raporları, https://www.teias.gov.tr/kurulu-guc-raporlari

[5] Padmavathi K, Daniel SA. Performance analysis of a 3MWp grid connected solar photovoltaic power plant in India. Energy for Sustainable Development, 2013;17:615-25.

[6] Eke R, Demircan H. Performance analysis of a multi crystalline Si photovoltaic module under Muğla climatic conditions. Energy Conversion and Management, 2013;65:580-86.

[7] Başoğlu ME, Kazdaloğlu A, Erfidan T, Bilgin MZ, Çakır B. Performance analyzes of different photovoltaic module technologies under İzmit, Kocaeli climatic conditions. Renewable and Sustainable Energy Reviews, 2015;52:357-65.

[8] Berwal AK, Kumar S, Kumari N, Kumar V, Hlaeem A. Design and analysis of rooftop grid tied 50kW capacity solar photovoltaic (SPV) power plant. Renewable and Sustainable Energy Reviews, 2017;77:1288-99.

[9] Yadav S, Bajpai U. Performance evaluation of a rooftop solar photovoltaic power plant in Northern India. Energy for Sustainable Development, 2018;43:130-38.

[10] Martin-Martinez S, Canas-Carreton M, Honrubia-Escribano, Gomez-Lazaro E, Performance evaluation of large solar photovoltaic power plants in Spain. Energy Conversion and Management, 2019;183:515-28.

[11] Kumar M, Chandel SS, Kumar A. Performance analysis of a 10MWp utility scale grid-connected canal-top photovoltaic power plant under Indian climatic conditions. Energy, 2020;117903.

[12] Al-Rasheedi M, Gueymard CA, Al-Khayat M, Ismail A, Lee JA, Al-Duaj H, Performance evaluation of a utility-scale dual-technology photovoltaic power plant at the Shagaya Renewable Energy Park in Kuwait. Renewable and Sustainable Energy Reviews, 2020:133;110139.

[13] Bentouba S, Bourouis M, Zioui N, Pirashanthan A, Velauthapillai D, Performance assessment of a 20 MW photovoltaic power plant in a hot climate using real data and simulation tools. Energy Reports, 2021;7:7297-314.

[14] Erdin C, Özkaya G. Turkey's 2023 energy strategies and investment opportunities for renewable energy sources: Site selection based on Electre. Sustainability, 2019;11:2136.

[15] Badel Kocagöz Demir, Yüksek Lisans Tezi Fotovoltaik Güç Sistemlerinde Performansı Etkileyen Parametrelerin İncelenmesi: Bayburt Güneş Enerjisi Santrali, Tez No: 738871, 2022

[16] BP Statistical Review of World Energy, https://ourworldindata.org/grapher/global-energy-consumption-source?country=~OWID_WRL

[17] Ministry of Energy and Natural Resources in Turkey, https://gepa.enerji.gov.tr/MyCalculator/

[18] Ministry of Energy and Natural Resources in Turkey, https://enerji.gov.tr/bilgi-merkezi-enerji-elektrik

[19] Cañete C, Carretero J, Sidrach-de-Cardona M. Energy performance of different photovoltaic module technologies under outdoor conditions. Energy 2014;65:295-302.

[20] Carr AJ, Pryor T. A comparison of the performance of different PV module types in temperate climates. Solar Energy 2004;76:285-294.

[21] Leloux J, Narvarte L, Trebosc D. Review of the performance of residential PV systems in France. Renewable and Sustainable Energy Reviews 2012;16:1369-1376.

[22] Moore LM, Post HN, Five years of operating experience at a large. Utility scale photovoltaic generating plant. Progress in Photovoltaics: Research and Applications 2008;16:249-259.

A KA BAND MMIC BASED VOLTAGE CONTROLLED OSCILLATOR FOR MODERN COMMUNICATION INFRASTRUCTURES

Hüseyin Şerif SAVCI[1]

INTRODUCTION

The demand for higher speed and more data pushed modern cellular bands towards upper millimeter wave frequencies. The 5G recent deployments and 6G near-future plans are already covering K-band, Ka band and even V-band operations. These systems are being designed as Time Division Duplexing which support channel bandwidths from 50MHz up to 2000MHz dependent on the Sub-Carrier Spacing in use. Among several other components the frequency synthesizers are gained extreme importance in such communication systems [1]. Most of these wireless systems use heterodyne architecture which requires several mixers and, thus, signal generators in the form of millimeter wave Voltage Controlled Oscillators. Although CMOS is the technology of choice for many RF circuits due to its higher integration potential with rest of the system, it still is not optimum in terms of performance versus cost to meet the millimeter wave design requirements [2]. For this reason, GaAs-based Pseudo-Morphic High Electron Mobility Transistor technology remains to be the primary technology of choice owing to their very high maximum oscillation frequency and cost-advantageous nature [3]. Majority of the Ka-band up and down converters in the market has an integrated frequency doublers or frequency quadruplers which works with X-band VCO. This chapter summarizes the design steps of MMIC VCO and presents a Ka band MMIC VCO with designs details.

This chapter is organized as follows: Next section presents VCO design approaches in general and the one port design procedure is detailed using a sample MMIC VCO design. The design details and schematics of the sample VCO are presented thorough the section. After that the physical layout and the EM model simulation results of the sample VCO are given in the following section. Finally, a conclusion based on the MMIC VCO using one port approach is presented.

MMIC VCO DESIGN

Most of the modern MMIC VCO design uses two different approaches. These are the Two port approach namely Feedback based Approach and one port approach namely Negative Resistance based Approach. The primary difference between them is the design methodology.

1   Hüseyin Şerif Savcı, Istanbul Medipol University, College of Engineering and Natural Sciences, Department of Electrical and Electronics Engineering, Istanbul, TURKEY.

In the two-port approach the oscillators are realized according to system block diagram shown in Figure 1.
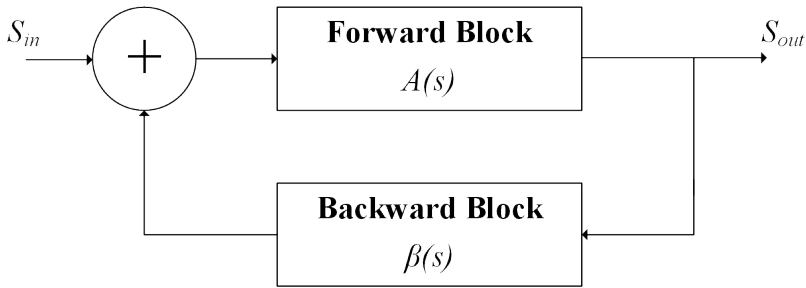
Figure 1. Two-port Approach.

The system functional for this block diagram is given in Equation 1.

$$H = \frac{S_{out}}{S_{in}} = \left[\frac{A(s)}{1 - A(s)\beta(s)} \right] \tag{1}$$

Here the oscillation starts when the system in Figure 1 satisfies the Barkhausen's criteria for oscillation. The Barkhausen's criteria is given in Equation 2 and Equation 3 as follows.

$$\|A(s)\beta(s)\| = 1 \tag{2}$$

$$\angle A(s)\beta(s) = 360^{o} \tag{3}$$

In the one port approach the system is in cascaded form rather than in a feedback configuration. Figure 2 depicts the one port approach.
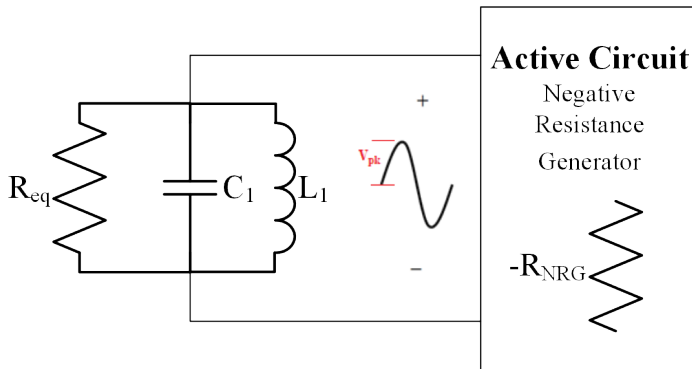
Figure 2. One-port Approach.

The resonator circuit is an inductor-capacitor, LC, tank network which determines the oscillation frequency. It has a finite equivalent resistance, R_{eq}. When supplied with initial energy the LC tank would oscillate as long as this equivalent resistance is compensated with a negative resistance, $-R_{NGR}$ generated by the active circuit. The amplitude of oscillation would depend on the initial energy injected into the tank.

Therefore, amount of negative resistance plays a critical role for the continuation of the oscillation. Equations 4 and 5 show the amount of energy and oscillation frequency for an ideal tank circuit consisting of parallel L1 and C1 components.

$$Energy = \frac{1}{2}C_1V_{pk}^2 \qquad (4)$$

$$Frequency = \frac{1}{2\pi\sqrt{C_1L_1}} \qquad (5)$$

Even though the initial energy supplied by the active circuit starts the oscillation, it does not guarantee a sustained one. If the resistive loss associated with the L and C components is greater than the initial energy the oscillation will eventually die. In other words, if the negative resistance generated by the active circuit is smaller than the tank resistance, the energy will decay over time and the oscillation will diminish after its initial start. On the other hand, a negative resistance much greater than the tank resistance means the excessive energy supplied by the active circuit. The oscillation keeps growing. This may eventually result in failure of the circuitry and stop of the oscillation. The amount of energy injected into the tank should be equal to the energy lost in the tank circuit. Energy injection to the LC resonator would be similar to cancelling out the tank resistance with the negative resistance. A steady oscillation only happens when the negative resistance is equal to the tank equivalent resistance. In this case oscillation will start with an initial energy and once the steady state is reached it will be sustained. This means the initial negative resistance generated by the active circuit should change over time and settle to the equivalent tank resistance, R_{eq}.

The oscillator generates a periodic signal at its output. There will be a series of harmonics at the output as well. In the real life, the output of oscillator is not a perfectly frequency stable. The output will fluctuate as a function of time so the instantaneous values of magnitude and frequency would be changing. This concept can be modeled with the Equation 6.

$$V_{out}(t) = (V_o + n_{noise}(t)\ \cos(\omega t + \theta + \theta_{noise}(t) \qquad (6)$$

The random fluctuations in time domain will result in a skirt on both side of the frequency response. The output is smeared out in the frequency domain. This smearing is mostly due to the phase variation. This it is known as Phase Noise of the oscillator. Oscillators phase noise spectrum is modeled with Leeson's equation [1] which is largely dependent on the inverse of loaded-quality factor of the resonator. Leeson's single-sideband phase noise expression in dBc/Hz is given in Equaiton 7. The spectrum of noise is modeled for an offset frequency of f_m.

$$L(f_m) = 10 \log \left[\frac{1}{2} \left(\left(\frac{f_o}{2Q_L f_m} \right)^2 + 1 \right) \left(\frac{f_c}{f_m} + 1 \right) \left(\frac{FkT}{P_s} \right) \right] \qquad (7)$$

Here, f_o is the oscillating frequency, f_c is the corner frequency of the flicker noise, P_s is the signal level at the oscillator active circuit input given as the available power and Q_L is the loaded quality factor of the resonator. The more the loaded quality factor of the resonator, the less the phase noise oscillator would have. Furthermore, the power supply and control voltage supply used in the oscillator should be highly stable as any supply noise will directly modulate oscillator's output.

An example of one port approach will be given in from of MMIC VCO in a 0.15 um InGaAs pHEMT process with 2 mils substrate thickness and depletion mode devices. Since pHEMT devices are depletion mode transistors, the feedback approach is hard to implement due to the mismatch of DC voltage level [4][5]. Thus, most MMIC designs adopt the negative resistance approach for its simplicity. Figure 3 represents the block diagram of the VCO with a negative resistance approach. A block diagram consists of a resonator, quarter wavelength transmission line, negative resistance generator, termination network, and coupled line filter. The resonator determines the oscillation frequency and fulfills the phase condition for oscillation startup. The negative resistance generator is the gain section that consists of a common source HEMT with series capacitive feedback. This section sources the necessary power via a negative resistance that drives the oscillation. The termination network and coupled line filter set a limit to delivered power and the phase noise of oscillation. Finally, the topology isolates the frequency-determining resonator from the load changes. The detailed design steps are explained in [8].
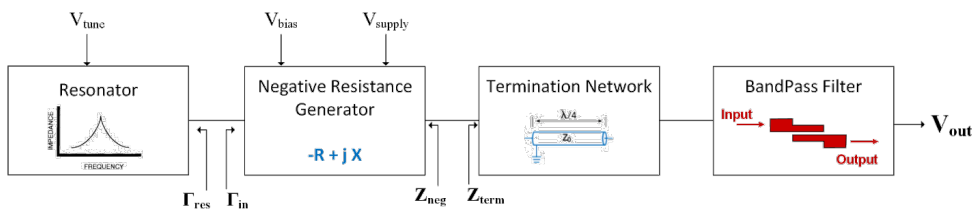
Figure 3. VCO Block Diagram with Negative Resistance Approach.

Figures 4 shows the resonator circuit. The drain-source connected device M1 is the voltage controlled variable capacitor which is a 100μm pHEMT device. The inductors L1 and L2 are used as RF chokes. Capacitors C1 and C2 are used for DC noise suppression.
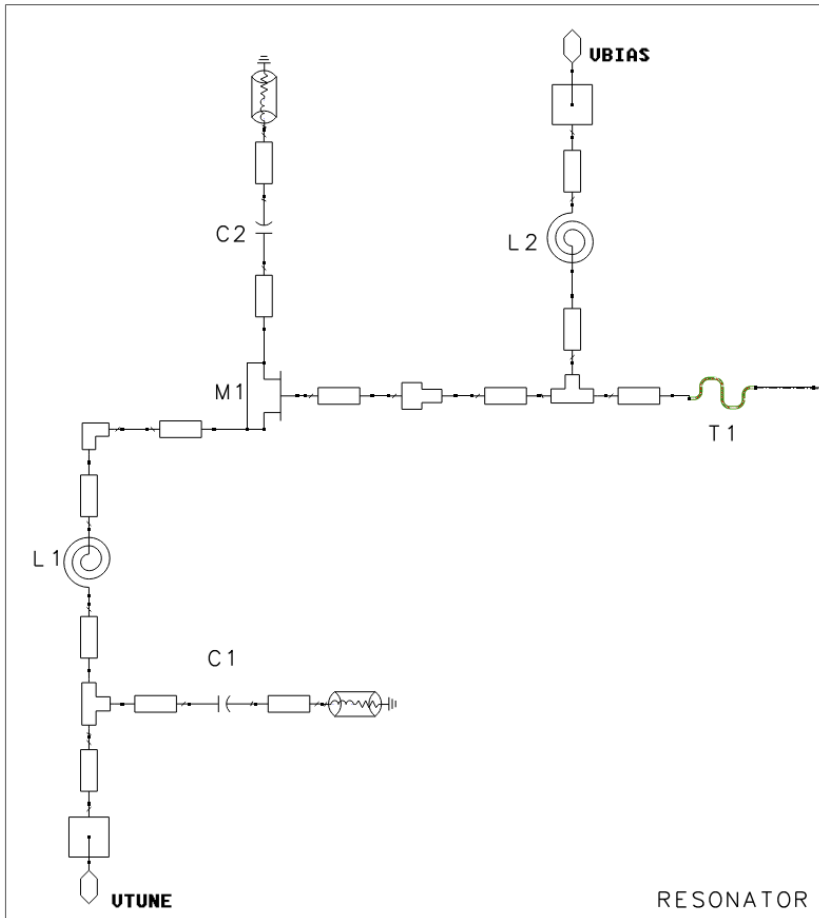
Figure 4. Resonator Network

Figures 5 shows the negative resistance generator network. The oscillation occurs when the pHEMT device is out of the stability region. Therefore, the design starts with destabilizing the transistor. The source feedback network ensures instability. A resistor or capacitor can be used for that purpose. In this design, a capacitor with an inductor is used. The core of negative resistance generator is realized with a 100μm pHEMT transistor M2 with source feedback topology consisting of inductor L5 and capacitor C5. The C5 is realized as series connected two capacitors, C5.1 and C5.2 to achieve small enough capacitance with available MiM capacitors from the process design kit. Inductors L3, L4, and capacitor, C3 are part of the supply network for RF choke and supply bypass reasons. Since inductor L4 is directly connected to the drain of the transistor, it has been chosen to be relatively smaller than the inductor L3. The feedback network consisting of L5 and C5 is applied to the pHEMT and adjusted to maximize the negative resistance at node Term looking into the circuit, which is Zneg in Figure 3.
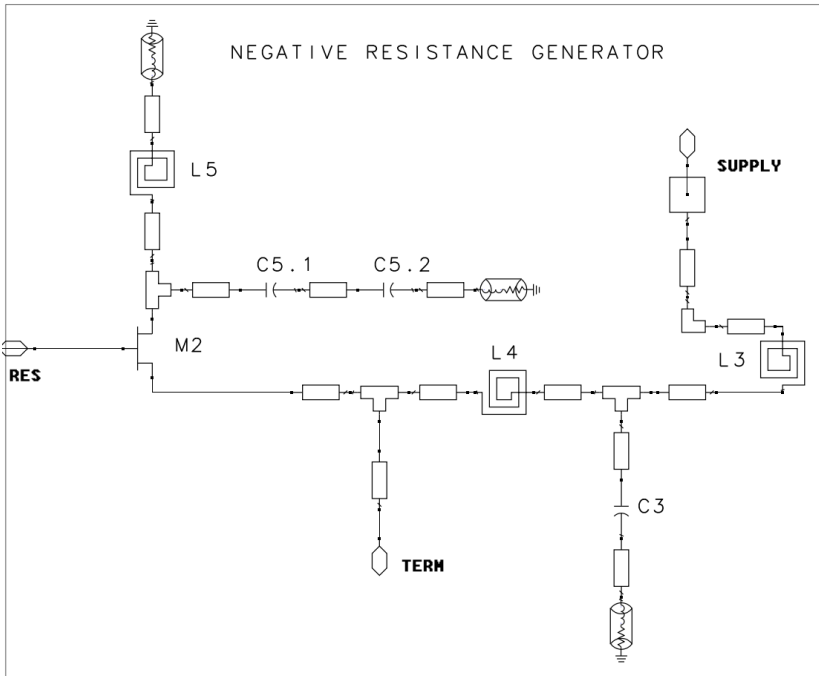
Figure 5. Negative Resistance Generator.

A proper DC bias point must be obtained for the negative resistance network. Figure 6 shows the obtained gate, drain, and source voltages and relative Vbias-Idrain curve.
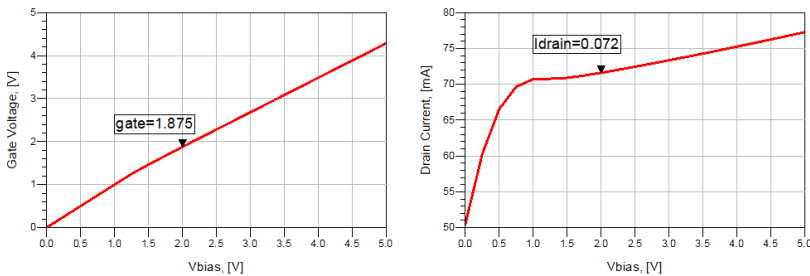
Figure 6. DC Bias Points of Negative Resistance Generator.

The figure on the left shows that there is a 125mV IR drop between Vbias and the voltage at the gate due to gate leakage current. The figure on the right shows that the transistor leaves the linear region approximately after 1.5 V of bias voltage. Therefore, choosing Vbias as 2 V is significantly improving the transconductance, gm, due to moderate inversion effects [2].

For the oscillation to startup, the absolute value of the reflection coefficient of Γres and the input impedance of the termination network should be unstable, meeting the condition in Equation 8. Also, the QWT helps to satisfy the phase condition in Equation 9.

$$| \Gamma_{res} | = \frac{1}{| \Gamma_{in} |} \tag{8}$$

$$\arg(\Gamma_{res}) = \arg\left(\frac{1}{\Gamma_{in}}\right) \tag{9}$$

The length of the quarter wavelength transmission line was obtained as 900um. Figure 7 shows layouts of two different forms of the same quarter wavelength transmission line. The meandered line on the left and curved line on the right. Figure 8 shows the reflection coefficient on the smith chart when the lines are short-circuited. The open circuit observation on the smith chart verifies the line was indeed quarter wavelength long.
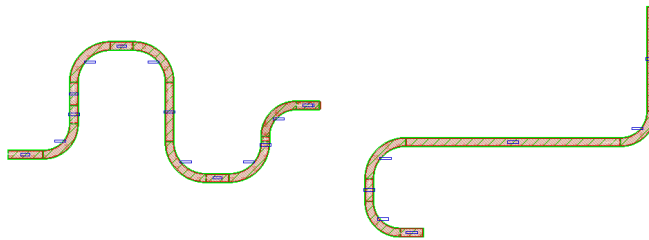
Figure 7. Layout Views of the same Quarter Wavelength Transmission Lines at 30-32 GHz.
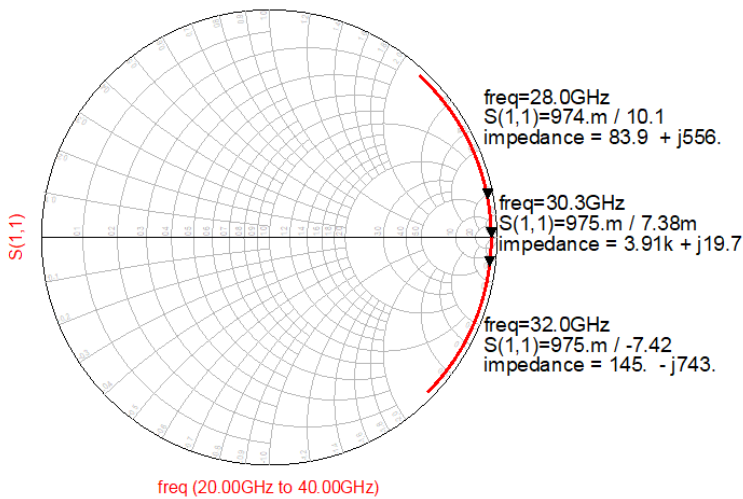
Figure 8. The Reflection Coefficient of Short-Circuited Quarter-Wave Transmission Line.

Figure 9 shows the final stage of the VCO with quarter-wave long transmission line as a termination network integrated with the bandpass filter and output matching. A shunt inductor is added to the output of a coupled line for impedance matching. They play an important role to maximize the transistor's input impedance. Here the bandpass filter is used to reject any out-of-band signals. The VCO design at millimeter wave frequencies includes number of distributed component due to the smaller sizes.

Equation 8 defines the requirement for the termination network to transfer the load impedance RL to Zterm as seen in Figure 3. This requirement is critical for a quick start and steady continuation of oscillation. The termination network and bandpass filter satisfy the condition in Equation 10.

$$Z_{term} = R_{term} + jX_{term} = -\frac{R_{neg}}{3} - jX_{neg} \tag{10}$$
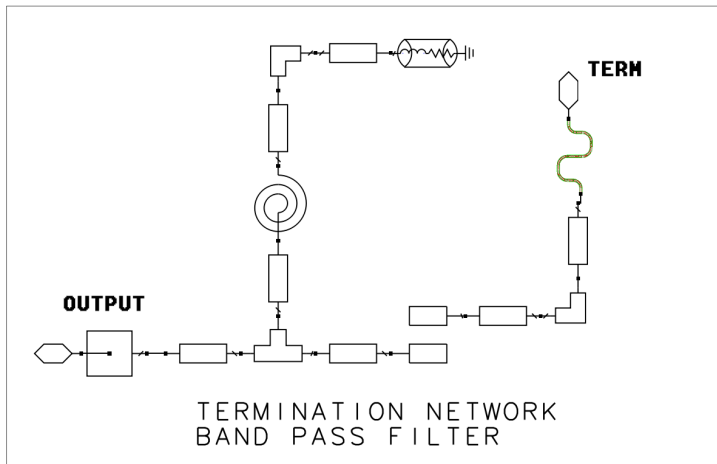
Figure 9. Termination network and coupled line bandpass filter.

Figure 10 shows the schematic-level simulation results of the circuit in Figure 3. It is seen in Figure 10 that the designed VCO exhibits an oscillation of 5.3V peak to peak at the termination output. The coupled line filter is revised to reduce the insertion loss. The voltage swing at the output of the VCO is 700mV peak to peak. Although the spacing of the coupled line is 4um which is the process minimum, the coupled line presents excessive insertion loss in the passband. Further optimization of the filter would improve its passband performance. The millimeter wave oscillators generally suffer from low power at the output. An amplifier as the output stage is generally used in designs to boost the output signal and isolate the circuit from varying load conditions.
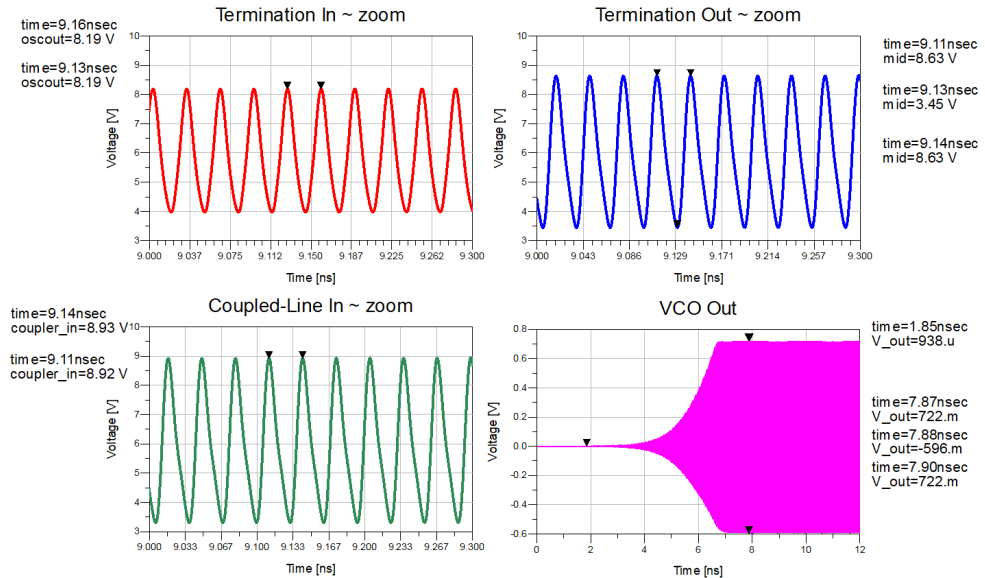
Figure 10. Schematic Level Simulation Results of the VCO.

SIMULATIONS AND EM MODEL

Figure 11 shows the layout of the designed VCO. EM model is generated by numerically Maxwell's equations to find the multi-port scattering response of this structure. Keysight's ADS Momentum software is used. Here 100 cells per wavelength are used for finer mesh generation to improve the accuracy of the solution. The co-simulations are done using the foundry-supplied pHEMT model on the EM solution.

The total die area is 1000um x 1600um, including scribe lines. Multiple Through-Wafer Via structures are used to ensure good grounding. The pad diagram of the chip is described in clock-wise manner starting from the lower left corner as follows: The control voltage pad is placed on the left. After that, the bias pad of the negative resistance generator is seen on the top left. Two ground pads come after the bias pad on the top middle. The down-bonds on these pads are to make sure that the chip gets proper DC ground in case the conductivity of epoxy is questioned. Right after the second ground pad, the power supply pad is placed. On the right output of VCO is routed to a GSG (Ground/Signal/Ground) pad. Such a pad placement eases the probe access to the die as no other probe will be close by the GSG RF probe.
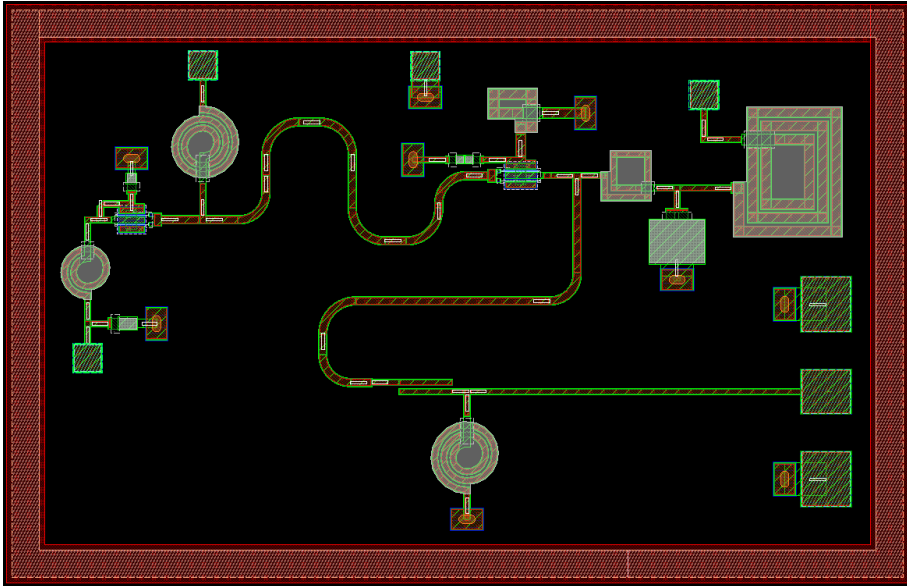
Figure 11. Layout View of the Designed VCO.

The bias point for the negative impedance generation can be optimized based on the EM co-simulations. It is observed that the output voltage is slightly reduced for Vcontrol of 2V. But VCO is still in tunable range between 2V to 6 V. The bias voltage determines the optimum control voltage range for which the output amplitude varies. As the control voltage decreases, the strength of oscillation gets reduced. Figure 12 and Figure 13 show the EM co-simulation results of the designed circuit while having the control voltage of 3V and 6V, respectively.
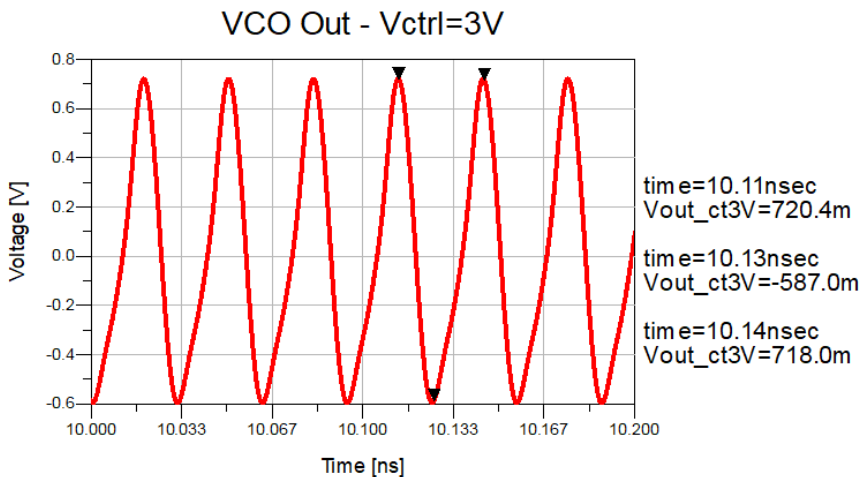
Figure 12. Output of the Designed VCO with VCTRL=3V.
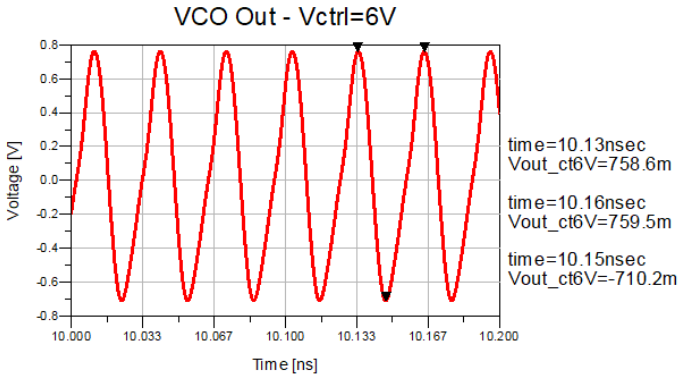
Figure 13. Output of the Designed VCO with VCTRL=6V.

Finally, Figure 14 shows the designed VCO output while sweeping the control voltage between 2V-6V in both time domain and frequency domain.
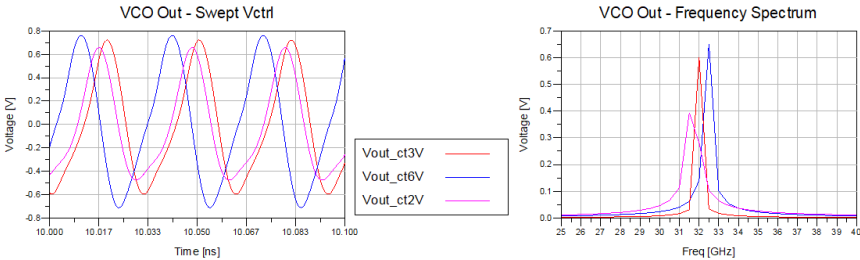
Figure 14. Time and Frequency response of VCO for Vctrl from 2V to 6V.

Table 1. Literature Comparison.

ref	Technology	Frequency (GHz)	Bandwidth (MHz)	Area (mm²)
[6]	InGaAs pHEMT	37.8	452	2.1
[7]	GaAs HEMT	30.5	2300	4.4
[9]	InGaP/GaAs HBT	38.5	1100	3.2
[10] *	InGaAs pHEMT	28.3	3800	3D (0.5)
[11]	InGaP/GaAs HBT	36	1900	3.76
[12]	GaAs pHEMT	34.5	3000	5
[13]	GaInP/GaAs HBT	34.2	400	1.56
This Work	InGaAs pHEMT	32	1200	1.6

CONCLUSION

In this chapter one port-based design approach using negative resistance generation is explained on a fully integrated MMIC VCO sample operating in the millimeter wave bands. The VCO shows a successful sustained oscillation at the target frequency where output reflection coefficient was above unity. An oscillation frequency tuning range of 31.4-32.6 GHz, with an output amplitude of 700mV, was achieved while consuming 100 mA current from a 6 V supply voltage.

REFERENCES

[1]. Hueber, G., and Niknejad A.M., Millimeter-Wave Circuits for 5G and Radar, The Cambridge RF and Microwave Engineering Series, Cambridge, 2019.

[2]. Razavi, B., RF Microelectronics, 2nd ed., Prentice Hall Communications Engineering and Emerging Technologies Series, Upper Saddle River, NJ, 2011.

[3]. Chen, P.Y., Tang, Y.L., Wang, H., Wang, Y.C., Chao P.C. and Chen, C.H., "A 39-46 GHz MMIC HBT triple-push VCO using cascode configuration", Proceedings of IEEE Asia-Pacific Conference on ASIC, 2002, (61-64).

[4]. Marsh, S., Practical MMIC Design, Artech House, Norwood, MA, 2006

[5]. Cheng, K.K.M., Oscillators: RFIC and MMIC Design and Technology, IET Circuits, Devices and Systems Series 13, The Institution of Engineering and Technology, London, 2009.

[6]. Chang, C.L., Tseng, C.H. and Chang, H.Y., "A New Monolithic Ka-Band Filter-Based Voltage-Controlled Oscillator Using 0.15 μm GaAs pHEMT Technology," IEEE Microwave and Wireless Components Letters, Feb. 2014, 24(2): 111-113.

[7]. Sevimli, O., Archer, J.W., and Griffiths, G.J., "GaAs HEMT monolithic voltage-controlled oscillators at 20 and 30 GHz incorporating Schottky varactor frequency tuning," IEEE Transaction on Microwave Theory and Techniques, 1998, 46(10):1572–1576.

[8]. Wilson, P.G. and Carver, R.D., "An easy-to-use FET DRO design procedure suited to most CAD programs," IEEE MTT-S International Microwave Symposium Digest, 1989, 3: 1033-1036.

[9]. Choumei, K., Matsuzuka, T., Suzuki, S., Hamano, S., Kawakami, K., Ogawa, N., Komaru, M. and Matsuda, Y., "A Ka-band direct oscillation HBT VCO MMIC with a parallel negative resistor circuit," IEEE MTT-S International Microwave Symposium Digest, 2005, 1175-1178.

[10]. Piernas, B., Nishikawa, K., Nakagawa, T., and Araki, K., "A compact and low-phase-noise Ka-band pHEMT-based VCO," IEEE Transactions on Microwave Theory and Techniques, 2003, 51(3): 778–783.

[11]. Mizutani, H., Nishida, K., Tsuru, M., Kawakami, K., Hieda, M., Taniguchi, E., Shimozawa, M, and Hirano, Y., "A millimeter-wave reflection-type dual-frequency VCO MMIC with a coupled line," 2010 IEEE MTT-S International Microwave Symposium, 2010, 288-291.

[12]. Ryu, S., Kim, H., Yim, J., Im, K., Kim, Y., Han, S., and Kim, B., "Ka-band PHEMT MMIC VCO with wide tuning range," Microwave and Optical Technology Letters, 2003, 39(4): 333–336.

[13]. Hilsenbeck, J., Lenk, F., Heinrich, W., and Wurfl, J., "Low phase noise MMIC VCOs for Ka-band applications with improved GaInP/GaAs-HBT technology," IEEE Gallium Arsenide Integrated Circuit Symposium, 2003, (223-226).

FEATURE SELECTION BASED ON GABOR FILTER AND BSO FOR DETECTING PARKINSON'S DISEASE

Pouya BOLOURCHI[1], Mohammadreza GHOLAMI[2]

INTRODUCTION

Parkinson's disease (PD) is one of the most common neurodegenerative disorders. PD is caused by a lack of a chemical messenger in the brain called dopamine which leads to impairment and deterioration of neurons and several brain regions. Due to the high number of recorded deaths and PD subjects per year, PD has become a concern in the neuroscience community. Predicting and early-stage diagnosis of PD are critical for medical issues. The progress of PD can be lowered remarkably by applying dopaminergic medications in the early years of the disease. Different modalities are used for PD, including single photon emission computed tomography (SPECT) [1] [2], positron emission tomography (PET) [3], magnetic resonance, and imaging (MRI) [4] [5] [6]. Structural magnetic resonance imaging (sMRI) is a neuroimaging technique in which the anatomy and pathology of the brain are examined. sMRI can be applied to quantify geometric structural properties such as the size and volume of a given structure or a cortical area's thickness. It is widely used to detect PD from HC due to its non-invasive and low-cost method. In this paper, we used the sMRI modality as input data.

On the other hand, computer-aided diagnostic (CAD) tools have recently been applied widely to related PD studies. Among these tools, machine learning applications have attracted the attention of scientists due to their high ability to identify diseases with high accuracy [7] [8] [9]. The process of machine learning methods generally consists of the following steps: In the first step, a set of images from a predefined modality is considered as the input data. Due to a lack of data or un-normalized data, a pre-processing step is applied to input data. The diffusion analysis, removing non-brain tissue, segmentation, and various interface mask-based methods are commonly used in the pre-processing stage [10] [11]. In the next step, feature extraction is applied to extract the most useful features. Some relevant feature extraction techniques for PD are local binary pattern (LBP), local gradient pattern (LGP), local neighbor descriptive pattern (LNDP), local neighbor gradient pattern (LNGP), different transformers, such as wavelet-based transformers, including complex and discrete wavelet, transform (CWT-DWT), Fourier transforms (FT),

1   Pouya Bolourchi, Assist. Prof. Dr., Final International University, https://orcid.org/0000-0003-3492-0617
2   Mohammadreza Gholami, Mr. Near East University

contourlet transforms (COT), Shearlet transforms (ST), and moment based methods including Zernike moment (ZM) and Chebyshev moment (CM) [12] [13] [14] [15] [16] [17] [18]. All the above-mentioned methods have their advantages and drawbacks. For instance, wavelet transformers are easy to apply but are associated with low accuracy. Moments-based methods have shown their ability to achieve better performance and higher accuracy; however, their significant limitation is normalization's complexity. FT, CWT-DWT, and COT suffer from the fact that they only operate in the frequency domain. To overcome this problem, we deploy the Gabor filter (GF), an improved FT, which can work in the time and frequency domain. GF is a combination of FT and Gaussian distribution function (GDF). Due to its high performance, we utilized GF in the feature extraction stage of our study.

The feature selection step is vital to reduce dimensionality by maintaining accuracy. The irrelevant features lead to an increase in the complexity of the problem and a decrease in the accuracy. As a result, some researchers tried to reduce the number of features to obtain higher efficiency and accuracy in detecting PD subjects. Due to the nonlinearity and nature of the problem, the feature selection can be considered an NP-hardness optimization problem that can be solved using meta-heuristic algorithms. The genetic algorithm (GA), particle swarm optimization (PSO), ant colony optimization, artificial bee colony, and brainstorm optimization (BSO) have shown their ability to solve such problems with high performance [19] [20] [21] [22] [23]. In this paper, we utilized a genetic BSO to select the most effective and ignore irrelevant features among all extracted features by the Gabor filter.

The final task in the ML procedure is classification. Many well-known classification methods have been used frequently in PD studies, such as k- nearest neighbor(k-NN), decision trees (DT), random forest (RF), neural network (NN), and support vector machine (SVM). SVM classifier has been used widely and achieved high performance in PD studies. This work uses SVM as a classifier for PD, healthy control (HC), and scans without evidence for dopaminergic deficit (SWEDD) subjects.

The rest of the paper is organized as follows. Section II presents the proposed method, including a description of the dataset, the principles, and the mathematical model of GF and BSO. Section III involves the results of applying the proposed BSO-GF-SVM method to the PD dataset. Also, the results are discussed in this section. Finally, section IV includes the conclusion of this study.

THE PROPOSED METHOD

In this section, we detailed all steps of the proposed ML-based method. First, the information regarding input data is given. Then a mathematical model of GF is provided. After that, a description of the performance of BSO and related equations are presented in this section. The block diagram of the proposed method is shown

in Fig.1. The input data involves sMRI modality images. All images are passed to the Gabor filter to extract the most relevant information. The extraction information includes both amplitude and phase information. Since the dimensionality of extracted features is exceptionally high, the feature selection technique must be applied to reduce the computational cost. In this regard, BSO is utilized for feature reduction. Finally, SVM with a linear kernel with the 10-fold-cross-validation technique is used for evaluating performances, including accuracy, sensitivity, and specificity.
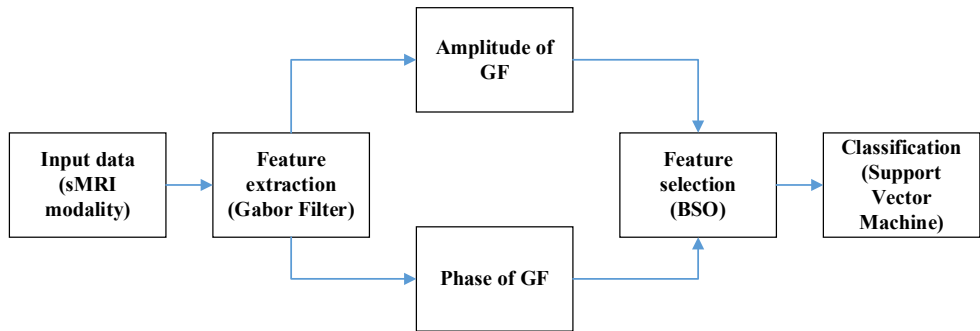
Figure 1. Schematic of the proposed method

Datasets

The 3.0T sMRI images utilized in this study are available on the Parkinson's progression markers initiative (PPMI) database [24]. The input data used in this study consists of a total of 238 subjects, including 62 normal control (NC), 142 PD, and 34 SWEDD subjects. Table I. gives some information regarding age, gender (male and female), weight, depression score, sleep score, olfaction score, and cognition score. The data and scores are states based on the mean and deviation values.

Gabor Filter

The Gabor filter is a linear filter that can be used in many applications, such as edge detection, segmentation, texture analysis, and feature extraction. As mentioned earlier, a GF is a combination of FT and GDF in which GDF acts as a kernel to measure the multiplication of FT [7]. GDF associated with the different frequency occurrences achieves the information related to the time domain. Eq. 1 and 2 show the FT and GDF equations, respectively. Also, the mathematical model of GF is given in Eq.3.

Table 1. Datasets

Subjects	NC	PD	SWEDD
No.	62	142	34
Age	61.0±10.7	61.5±9.4	61.7±9.8
Gender (male/female)	23/39	51/91	14/20
Weight	76.9±15.4	82.6±16.6	83.1±12.5

Depression score	5.1±1.2	5.3±1.5	5.6±1.3
Sleep score	33.4.2±4.7	22.5±8.5	29.9±8.4
Cognition score	28.2±1.1	27.5±2.1	26.9±2.9

$$\hat{f}(w) = \int_{-\infty}^{\infty} f(x)e^{-iwx}dx \tag{1}$$

$$g_a(t) = e^{-(t-\tau)^2/a^2} \tag{2}$$

$$G(f)(t,w) = \int_{-\infty}^{\infty} f(\tau)e^{-iw\tau}g_a(t-\tau)d\tau \tag{3}$$

where, t and ω are the time and frequency domains. τ represents the center and the spread of the window in GDF.

Both amplitude and phase information of the 2D-GF output of the sMRI images are used to extract the features. Figure 2. represents a sample of the PD brain as an input image and the amplitude and phase of GF as the output.
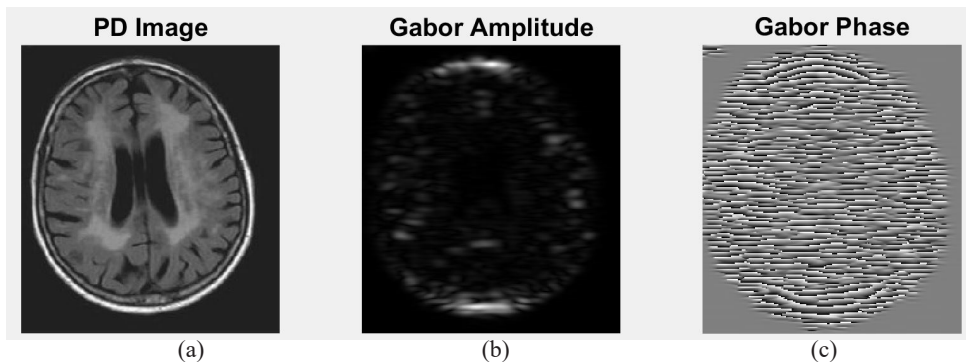
Figure 2. Applying GF on a PD sample: a. Original PD image b. GF amplitude c.GF phase

In this study, we applied a Gabor mask filter with a size of 25×25, which means that for amplitude, we extracted 625 features. Another 625 features are extracted for the phase information, which makes a total of 1250 features.

Brainstorm Optimization

The BSO algorithm is categorized as a swarm-based optimization algorithm and is widely used to solve nonlinear and multi-objective optimization problems. Generally, a solution to the problem is presented as an idea in BSO. The population comprises a group of ideas with different characteristics and solutions to the fitness function [23]. Like other optimization-based algorithms, the first generation, a set of ideas, is created randomly and updated after iterations based on the procedure of BSO. The following steps are done to update the ideas in iterations.

In the first step, similar ideas are grouped using a k-means clustering method. Based on the calculated fitness function, the best idea is selected as the center of

each cluster. To update the ideas and create new ideas (idea $_{new}$), the following tasks are applied to old ideas idea $_{old}$.

- Set the values to the pone−cluster, pone−center, and p $_{two}$−centers as BSO control parameters.

- A random number is considered for every idea in the generation. Then:

- If the random number is less than the pone−cluster, a cluster is selected randomly, and the cluster's center replaces the idea if the random number is less than the pone−center.

- If the random number is more significant than the pone−cluster, two clusters are selected randomly, and the chosen idea is updated towards the centers of selected clusters as Eq.4.

- Finally, the new idea is updated based on Eq.5 and 6 using a step-size parameter ξ and Gaussian distribution.

$$if\ rand < p_{two-centers}:$$

$$idea_{new_i} = r * center_{cr1} \pm (1 - r) * center_{cr2} \tag{4}$$

$$\delta = rand * \log \frac{(0.5 * \max\_iter - iter)}{k} \tag{5}$$

$$idea_{new_i} = idea_{new_i} \pm \delta * N(0,1) \tag{6}$$

Each idea in the BSO algorithm represents a possible status of extracted features. A population consists of a specific number of ideas. A new population is created by applying the BSO operations to the old population. As it shown in Table. 2, each idea with the size of a 1×1250 array. 1250 represents the number of extracted features from the images applying GF. Features represented with "one" are selected, and other features with "zero" are ignored. For example, in the given sample depicted in Table 2. features 1 and 1249 are chosen, and feature 2 and 1250 are ignored.

The accuracy is considered the fitness function.

$$ff_i = \max(acc_i) \tag{7}$$

where ff_i is the fitness function which should be maximized. acc_i is the accuracy of classifying related to idea i. The summarized information regarding the BSO parameters is given in Table 3.

Table 2. A sample ideas representing features statues

Feature	1	2	…	1249	1250
Particle value	1	1		1	0
Feature statues	selected	ignored		selected	ignored

Classification And Metrics

Based on statistical learning principles, the SVM is the most powerful classifier. The SVM algorithm has been successfully adapted to various applications in Gene application. During the training stage, SVM attempts to find the optimal class-separation hyperplane in the maximal margin, which is the distance between the support vectors on the boundary. Support vectors are located on two parallel hyperplanes, with a distance of 2/‖w‖ between them.

In [25],SVM was first used for pattern recognition. SVM takes into account a set of samples containing x (train set vector) and y (corresponding labeled set vector), and the decision surface is defined as:

$$f(x) = \sum_i a_i y_i K(x_i, x_i') + b \tag{8}$$

where a_i is the weight constant, which is optimized during the training process. K(.,.) is the kernel function, x_i are support vectors, b is bias and $y_i = \pm 1$ is the label of the sample x_i . K(.,.) describes the behavior of support vectors as the kernel function.

Table 3. BSO parameters

Parameter	Values
Maximum iterations	10000
Total idea size in each population (n)	100
Number of clusters (m)	4
$p_{one-cluster}$	0.8
$p_{one-centers}$	0.4
$p_{two-centers}$	0.5

In this study, SVM was used for classification with a linear kernel. A 10-fold cross-validation technique is used for training, validation, and testing. The performances applied in this study include sensitivity, specificity, and accuracy, which are defined as:

$$Sensitivity = TPR = \frac{TP}{TP + FN} \tag{9}$$

$$Specificity = TNR = \frac{TN}{TN + FP} \tag{10}$$

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

where TP, FN, TN, and FP denote true positive, false negative, true negative, and false-positive results respectively that are defined in the interval [0, 1].

RESULTS AND DISCUSSION

In this section, the proposed method is applied to the dataset given in Table. 1. First, the GF is used to extract the features from input images. The extracted features are fed into the SVM for classification. Then, the BSO is utilized to reduce the number of features, and the selected features are considered as input data of SVM. The results of the classification of PD/NC, SWEDD/NC, and PD/SWEDD using both GF and GF+BSO are provided in Tables. 4, 5, and 6, respectively.

The results show that the BSO can achieve higher accuracy using lower features. In addition, the BSO can reduce the number of selected features to 137. These results reveal that irrelevant features still exist in the feature extraction step and are ignored in the feature selection step. The fewer features can reduce the computational cost and improve the performance of classification tasks. In addition, as we can observe from Tables IV to VI, the highest accuracy is achieved in AD/HC, 96.17%. The accuracy of detecting PD/SWEDD is remarkably lower than PD/NC and SWEDD/NC, because the number of similar features in this category is relatively higher than in the other two categories.

Table 4. The results of the SVM classifier to classify PD/NC subjects.

Method	Feat.	ACC.	SEN	SPE
GF	1250	85.34	79.32	82.46
GF+ BSO	513	82.56	76.36	80.72
	496	91.47	83.12	86.89
	137	96.17	89.18	89.57

Table 5. The results of the SVM classifier to classify SWEDD/NC subjects.

Method	Feat.	ACC.	SEN	SPE
GF	1250	89.96	86.38	78.05
GF+ BSO	513	86.83	82.18	81.06
	496	90.38	91.46	87.42
	137	94.57	93.26	89.74

Table 6. The results of the SVM classifier to classify PD/SWEDD subjects.

Method	Feat.	ACC.	SEN	SPE
GF	1250	71.03	68.54	59.32
GF+ BSO	513	69.57	65.82	61.02
	496	74.32	72.86	68.74
	137	78.57	80.14	78.45

In addition, we applied other well-known feature extraction methods, such as CWT, WT, CM, and ZM, to the same dataset. The extracted features are fed into the SVM as a classifier to compare with the results obtained by the proposed method. The results of the classification of PD/NC, SWEDD/NC, and PD/SWEDD using different feature extraction methods are given in Figures 3,4, and 5, respectively. The

results show the efficiency of GF compared to other feature extraction methods. The results of CM and ZMs, are comparable with the proposed method. The moment-based methods had better performance rather than wavelet-based methods. The fitness function can be replaced by desired performance meters to achieve higher sensitivity and specificity. Also, it is possible to define a multi-objective fitness function at the selection stage.
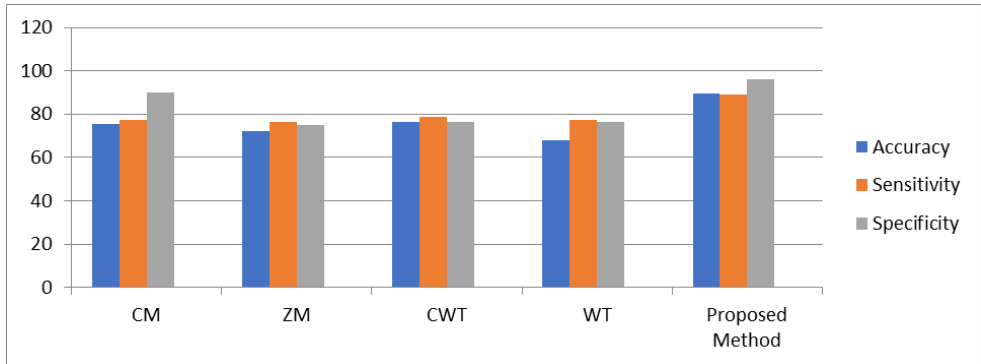
Figure 3. The performance comparison of feature extraction methods for both NC/PD
(Using SVM as a classifier)
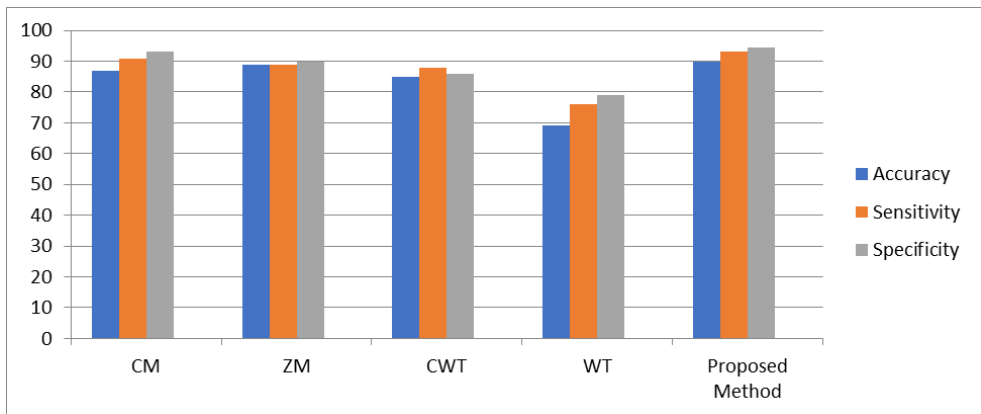
Figure 4. The performance comparison of feature extraction methods for both NC/SWEDD
(Using SVM as a classifier)
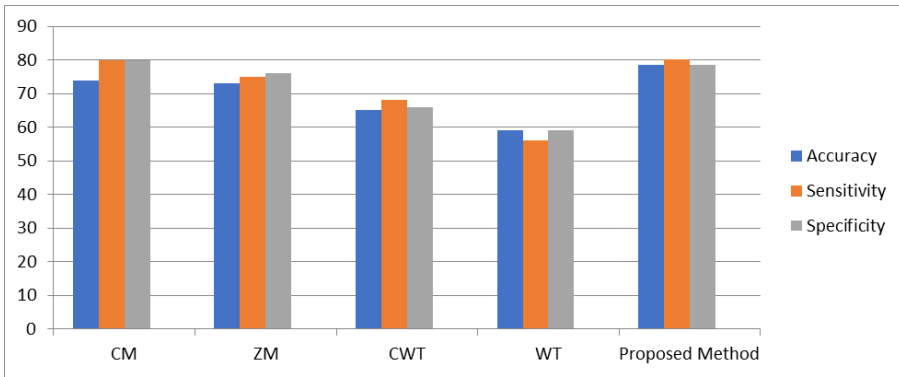
Figure 5. The performance comparison of feature extraction methods for both PD//SWEDD (Using SVM as a classifier)

As we expected, the overall results obtained by SVM are higher in terms of accuracy, sensitivity, and specificity. Among the other classifiers, the random forest has shown better efficiency in detecting PD subjects. Moreover, our results show that distinguishing the PD and SWEDD subjects can be done with higher accuracy than detecting PD against SWEDD subjects.

Table 7. Experimental results of the different classifiers (NC/PD classification).

Method	Accuracy	Sensitivity	Specificity
k-NN	80.14	83.27	79.85
Random Forest	94.13	90.24	83.16
Decision Tree	90.13	86.48	79.81
SVM	96.17	89.18	89.57

Table 8. Experimental results of the different classifiers (NC/SWEDD classification).

Method	Accuracy	Sensitivity	Specificity
kNN	80.01	82/46	76.39
Random Forest	90.48	89.37	80.16
Decision Tree	88.91	83.79	78.45
SVM	94.57	93.26	89.74

Table 9. Experimental results of the different classifiers (PD/SWEDD classification).

Method	Accuracy	Sensitivity	Specificity
kNN	63.18	68.66	61.48
Random Forest	75.56	79.32	70.16
Decision Tree	74.82	73.28	69.57
SVM	78.57	80.14	78.45

CONCLUSION

A novel approach using GF and BSO was proposed to distinguish the PD/HC, SWEDD/HC, and PD/SWEDD subjects. The sMRI images from 238 subjects, including 62 NC,142 PD, and 34 SWEDD, were selected as input data. A GF as an effective transformer was utilized to extract the features from the images. Then, a BSO algorithm was applied to reduce the number of features and computational costs and eliminate the irrelevant features. A standard 10-fold classification SVM with a linear kernel was used to classify the different groups using selected features. The results showed that GF-BSO achieved an accuracy of 96.17 percent by selecting only 137 features among 1250 features. The results were compared with well-known feature extraction methods and the results of the state-of-the-art papers which have used other classifiers. The high accuracy obtained in the classification task reveals the efficiency of the proposed approach in distinguishing PD, HC, and SWEDD subjects. Furthermore, fewer features make the proposed method applicable for medical scientists for better treatment performance.

REFERENCES

[1] H. Khachnaoui, N. Khlifa, and R. Mabrouk, "Machine Learning for Early Parkinson's Disease Identification within SWEDD Group Using Clinical and DaTSCAN SPECT Imaging Features," 2022.

[2] C. Y. Chien, S. W. Hsu, T. L. Lee, P. S. Sung, and C. C. Lin, "Using artificial neural network to discriminate parkinson's disease from other parkinsonisms by focusing on putamen of dopamine transporter SPECT images," Biomedicines, vol. 9, no. 1, pp. 1–11, 2021, doi: 10.3390/biomedicines9010012.

[3] E. Glaab et al., "Integrative analysis of blood metabolomics and PET brain neuroimaging data for Parkinson's disease," Neurobiol. Dis., vol. 124, no. December 2018, pp. 555–562, 2019, doi: 10.1016/j.nbd.2019.01.003.

[4] S. Chakraborty, S. Aich, and H. C. Kim, "Detection of Parkinson's disease from 3T T1 weighted MRI scans using 3D convolutional neural network," Diagnostics, vol. 10, no. 6, 2020, doi: 10.3390/diagnostics10060402.

[5] S. Sivaranjini and C. M. Sujatha, "Deep learning based diagnosis of Parkinson's disease using convolutional neural network," Multimed. Tools Appl., vol. 79, no. 21–22, pp. 15467–15479, 2020, doi: 10.1007/s11042-019-7469-8.

[6] A. S. Talai, Z. Ismail, J. Sedlacik, K. Boelmans, and N. D. Forkert, "Improved Automatic Morphology-Based Classification of Parkinson's Disease and Progressive Supranuclear Palsy," Clin. Neuroradiol., vol. 29, no. 4, pp. 605–614, 2019, doi: 10.1007/s00062-018-0727-8.

[7] H. W. Loh et al., "Gaborpdnet: Gabor transformation and deep neural network for parkinson's disease detection using eeg signals," Electron., vol. 10, no. 14, 2021, doi: 10.3390/electronics10141740.

[8] L. Shen et al., "Multiple Empirical Kernel Mapping Based Broad Learning System for Classification of Parkinson's Disease with Transcranial Sonography," Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS, vol. 2018-July, pp. 3132–3135, 2018, doi: 10.1109/EMBC.2018.8512990.

[9] B. Rana et al., "Regions-of-interest based automated diagnosis of Parkinson's disease using T1-weighted MRI," Expert Syst. Appl., vol. 42, no. 9, pp. 4506–4516, 2015, doi: 10.1016/j.eswa.2015.01.062.

[10] Y. Dai, Z. Tang, Y. Wang, and Z. Xu, "Data Driven Intelligent Diagnostics for Parkinson's Disease," IEEE Access, vol. 7, pp. 106941–106950, 2019, doi: 10.1109/ACCESS.2019.2931744.

[11] J. Xu et al., "A fully automatic framework for Parkinson's disease diagnosis by multi-modality images," Front. Neurosci., vol. 13, no. AUG, pp. 1–9, 2019, doi: 10.3389/fnins.2019.00874.

[12] E. S. A. El-Dahshan, T. Hosny, and A. B. M. Salem, "Hybrid intelligent techniques for MRI brain images classification," Digit. Signal Process. A Rev. J., vol. 20, no. 2, pp. 433–441, 2010, doi: 10.1016/j.dsp.2009.07.002.

[13] C. Zhang et al., "Dynamic Alterations of Spontaneous Neural Activity in Parkinson's Disease: A Resting-State fMRI Study," Front. Neurol., vol. 10, no. October, 2019, doi: 10.3389/fneur.2019.01052.

[14] Y. Wu et al., "Use of radiomic features and support vector machine to distinguish Parkinson's disease cases from normal controls," Ann. Transl. Med., vol. 7, no. 23, pp. 773–773, 2019, doi: 10.21037/atm.2019.11.26.

[15] S. Shinde et al., "Predictive markers for Parkinson's disease using deep neural nets on neuromelanin sensitive MRI," NeuroImage Clin., vol. 22, no. March, p. 101748, 2019, doi: 10.1016/j.nicl.2019.101748.

[16] R. Mukundan and S. Ong, "Discrete Orthognal Moment Features Using Chebyshev Polynomials," New Zeal. Int., no. ii, pp. 20–25, 2000, [Online]. Available: http://ir.canterbury.ac.nz/handle/10092/446

[17] C. H. Teh and R. T. Chin, "On image analysis by the methods of moments.," vol. I, no. 4, pp. 556–561, 1988, doi: 10.1109/cvpr.1988.196290.

[18] S. Ghosal, R. Mehrotra, and S. Louis, "ZERNIKE MOMENT-BASED FEATURE DETECTORS Center for Computational Math University of Colorado at Denver," pp. 934–938, 1994.

[19] T. M. Shami, A. A. El-Saleh, M. Alswaitti, Q. Al-Tashi, M. A. Summakieh, and S. Mirjalili, "Particle Swarm Optimization: A Comprehensive Survey," IEEE Access, vol. 10, pp. 10031–10061, 2022, doi: 10.1109/ACCESS.2022.3142859.

[20] K. Chandramouli and E. Izquierdo, "Image classification using chaotic particle swarm optimization," Proc. - Int. Conf. Image Process. ICIP, no. August, pp. 3001–3004, 2006, doi: 10.1109/ICIP.2006.312968.

[21] A. K. Kesarwani, M. Yadav, D. Singh, and G. D. Gautam, "A review on the recent applications of particle swarm optimization & genetic algorithm during antenna design," Mater. Today Proc., no. xxxx, pp. 2–4, 2022, doi: 10.1016/j.matpr.2022.02.200.

[22] Z. Cao, X. Hei, L. Wang, Y. Shi, and X. Rong, "An Improved Brain Storm Optimization with Differential Evolution Strategy for Applications of ANNs," Math. Probl. Eng., vol. 2015, 2015, doi: 10.1155/2015/923698.

[23] M. El-abd, "Global-best brain storm optimization algorithm," Swarm Evol. Comput., no. May, pp. 0–1, 2017, doi: 10.1016/j.swevo.2017.05.001.

[24] H. Lei et al., "Parkinson's Disease Diagnosis via Joint Learning from Multiple Modalities and Relations," IEEE J. Biomed. Heal. Informatics, vol. 23, no. 4, pp. 1437–1449, 2019, doi: 10.1109/JBHI.2018.2868420.

[25] J. Wu and C. Hicks, "Breast Cancer Type Classification Using Machine Learning," J. Pers. Med., vol. 11, no. 2, p. 61, Jan. 2021, doi: 10.3390/jpm11020061.

A REVIEW OF FPGA-BASED APPLICATIONS AND FPGA USAGE IN THE INDUSTRIAL AREA

Abdulkadir SADAY [1]

INTRODUCTION

Developments in industrial technologies also reveal many factors such as high speed, practicality, fast time-to-market, price-performance balance, and embedded systems. In today's industrial applications, the use of computer-aided systems and control systems managed with embedded hardware is becoming increasingly necessary due to digitalization. However, depending on consumption demands, the need to accelerate production and achieve more performance with less cost arises.

In industrial applications, it is necessary to reach high-performance control systems in a safe and flexible structure to catch up with the competition in the sector. In addition, the cost of production is an important factor for the products offered to the market. Reducing the cost, and shortening the time of the entry of the product into the market, depends on the performance of the technology used and the amount of energy consumed in high performance.

Today's consumer products and technologies are constantly updated and renewed depending on rapidly changing user demands. Although traditional control systems are sufficient to produce a solution for a need, restructuring them according to changing demands creates processes that require high costs and time. For this reason, manufacturers and developers have been in search of real-time, high-performance, and low-cost systems that can be configured in the field depending on the needs. In addition, with the increase of controlled systems in applications, the problem of synchronous control of more than one system has emerged. The limited nature of traditional processors has become a problem for multiple systems.

Field programmable gate arrays (FPGAs) are semiconductor digital circuits built on customizable logic blocks connected by programmable connections. FPGAs can be reprogrammed according to the desired application or functionality requirements. Due to their low power consumption and correspondingly high computational capacity, they are gaining increasing interest in industrial applications and academic studies. FPGAs are configured to execute a process as needed. Unlike traditional processors due to their architecture, they can be reconfigured in the field for the needs [1]. These structures have started to make FPGAs a choice for developers

1   Selcuk University, Faculty of Technology, Mechatronics Engineering, Konya, TURKIYE.

and manufacturers because of the cost-performance criterion. In current studies, FPGAs are used extensively, and their performance values are compared with other processor and microcontroller types. According to comparisons made in many areas such as signal processing, image processing, and artificial intelligence applications, it is shown that FPGAs offer up to 57 times faster speed with lower energy consumption than GPU and traditional processors [1, 2].

Despite their advantages, FPGAs are devices that take a long time to design due to their architecture, are designed with hardware description language (HDL), and require special hardware skills. Today, their not very intense industrial prevalence, limited knowledge and access to design resources make these devices a field that requires special expertise. One of the important obstacles to the widespread use of FPGAs in applications is the complexity of programming [3]. System-on-chip (SoC) platforms have been developed by combining FPGAs with embedded processors as a result of the work done by the developers to solve this problem. The purpose of these platforms is stated as to ensure that only the parts that need performance, not the entire application, are configured on the FPGA and to avoid design difficulties.

When current studies are examined [4-10], it is seen that FPGA and SoC hardware have started to come into focus in many different areas in industrial applications and research issues that require performance. Major sectors include industry, computing, aerospace, energy, defense, and medicine heavily. Experts state that FPGAs will see high demand in the coming years [11, 12]. Despite their high cost, still these devices are well suited for industrial applications due to their long lifetime, high bandwidth, functional security, flexible design support, and parallel processing capabilities. Therefore, the purpose of this article is to provide a review with an analysis of the use of FPGAs for industrial applications, primarily in the field of engineering.

FPGA DESCRIPTION

FPGAs are defined as reconfigurable logical blocks containing programmable logic gates, connected via interconnection networks. They can be reprogrammed depending on the need after production. With these features, FPGAs differ from application-specific integrated circuits (ASICs) produced for special functions. FPGAs are generally produced as SRAM-based reprogrammable circuits. Today, giant processor manufacturers such as Intel and AMD have included the first FPGA manufacturers such as Xilinx and Altera, and the development of these technologies has gained great momentum. FPGAs offer users the opportunity to reconfigure to meet their post-production needs. That's why they got the name programmable in the field. They are frequently used in commercial applications such as communications and image processing, where flexibility and parallel computing are needed. The FPGA architecture consists of a matrix of configurable logic blocks (CLB). This matrix is surrounded by up to 1000 input-output blocks (IOBs). All this structure provides communication between itself with an interconnection [13-16]. The representation of the structure is shown in Figure 1.
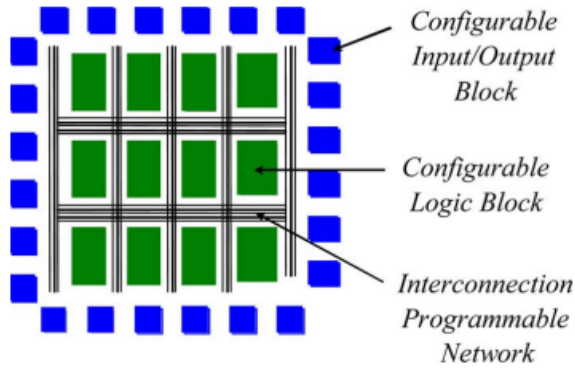
Figure 1. FPGA architecture [16]

With the developments in FPGA technology, embedded processor blocks such as RAM and DSP accelerators have been added to the architecture. This structure emerged as a complement to the original FPGA matrix. This hardware is defined with Verilog and VHDL languages depending on the hardware description language (HDL) [17, 18]. In the definition phase, four different steps must be followed to integrate the circuit. While design and definition processes are provided in the hardware language, progress is made by verifying each stage. The four steps of the identification phase are given:

1- Base system level where circuit properties are determined

2- Behavior level of the circuit

3- Register transfer level (RTL) at which component identification is made

4- Physical synthesis level depending on hardware features

Applying each level means defining the circuit to be created with HDL. If any of these definition levels for the FPGA is missing or incorrect, circuit integrity cannot be achieved. Therefore, at the design stage, each identification level must be successfully accomplished. The flow of the design approach is given in Figure 2.
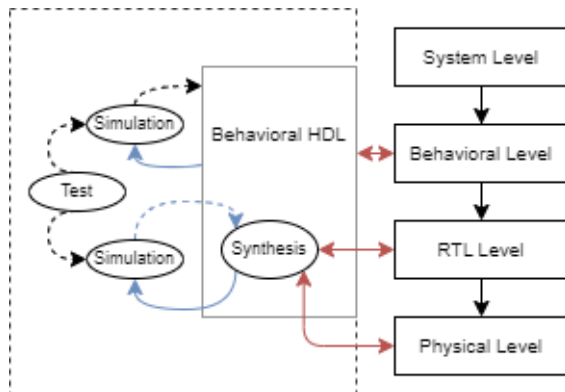
Figure 2. Design approach

After FPGAs are deployed in the field, they allow hardware upgrades to add extra features or fix bugs with a new bitstream as needed. This makes FPGAs a focal solution, especially for fast production cycles in the technology-dependent market. The level of reconfigurability allows the creation of application-specific hardware rather than a fixed general architecture such as CPUs and GPUs. In this way, FPGAs can be used in a hardware design that includes application-specific bandwidth or instruction set and can offer higher efficiency than CPU or GPU [19, 20].

Today, FPGAs are increasingly used in industry for complex functions and calculations. Creating an arithmetic logic unit (ALU) using single hardware, designing complete digital systems such as communication units, and configuring it as a control unit in remote control units such as satellite communication is a distribution that shows the potential of today's usage areas of FPGAs to become widespread.

INDUSTRIAL APPLICATIONS

In this section, industrial applications developed using FPGAs and the plans made by head manufacturers for the future of this hardware in recent years are analyzed. A review is provided by considering the real-time applications designed, control techniques, and areas of use. Due to their reconfigurable architecture, FPGAs have become suitable hardware for different industries and applications. With the latest developments, AMD Xilinx and Intel companies, which are the leading processor manufacturers in the industry, have started to offer FPGA devices, software, and ready-made IP package solutions specially developed for different industries. Within the framework of the market strategies and visions of the manufacturers in the industry, common application areas are gathered under the main prominent categories.

Aerospace, Space, and Defense

Because of their radiation resistance and reconfigurable properties, FPGAs are ideal equipment for important signal-processing applications such as image processing, SDR, and waveform generation. Aerospace and defense fields often place less emphasis on hardware cost for development than the regular consumer, making FPGAs more likely to be used by developers in this industry. Studies in this field generally include progressive radar systems, satellite communication systems, missile guidance systems, electronic warfare, and secure communication systems. These systems require real-time processing and calculation of large amounts of data. The phased array algorithm used for radars uses independent antenna element arrangement to manage the radiation pattern [21, 22]. Because this design requires high bandwidth along with high-performance digital signal processing (DSP), it can often be implemented using an FPGA solution. Communication security and low latency are needed for military and warfare systems [23, 24]. In these systems, using wireless encoders and decoders makes FPGAs with the ability to be reconfigured in the field a necessity.

Studies such as exposure estimation of spacecraft, moving target defense, development of wireless sensors for secure data transfer, fault prediction in a space radiation environment, and creation of high-capacity routers for satellite backbone networks in aerospace, space and defense sectors [25-28] It benefits from data processing capacity and functional security features.
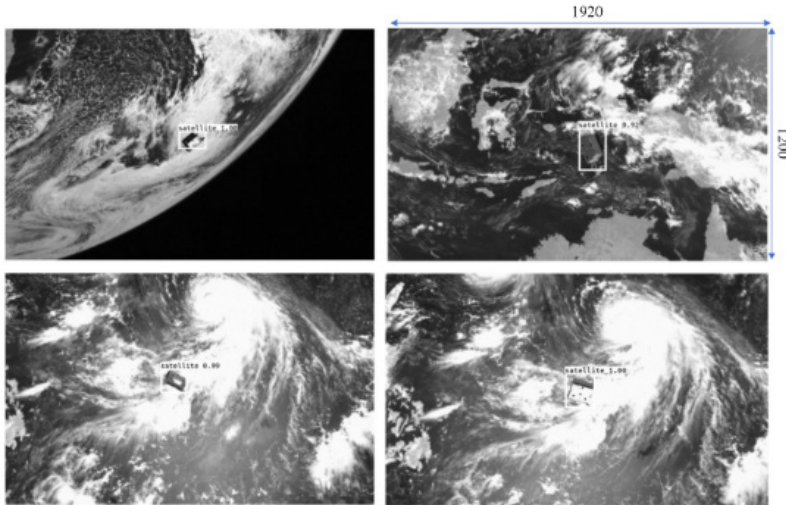

Figure 3. Spacecraft detection and bounding [25]

Automotive

In the automotive sector, many systems such as driving assistance systems, comfort, and safety, and in-car information systems are offered as package services today. Autonomous driving and driving assistance systems require a high level of safety to adapt to rapid mobility. FPGAs used in the automotive industry offer advantages with low latency in terms of security in functions that require high performance. Various sensors in automotive technologies generate large amounts of data that must be collected in real-time under a single system [29]. With the high bandwidth and computation speed they provide, FPGAs are critically helpful in collecting, preprocessing, and distributing this data, combining it with various software algorithms including artificial intelligence, and creating a safe and efficient application for vehicle control [30]. Studies with FPGA in the automotive field [31, 32] have been developed to obtain a real-time response in all operations from the vehicle's internal control system to the battery management system. The use of FPGAs in the design phase provides flexibility and cost savings during the development and testing of various camera and data transfer structures. Due to the widespread use of smart vehicles, the limitations of the fixed hardware circuits in the existing electronic control units have led the developers to seek new technologies. Flexible hardware upgrades and reconfiguration make FPGAs a viable strategy for these investigations. In recent studies [33, 34] remote reconfigurable ECU systems

are recommended for automotive systems. This feature allows manufacturers to remotely perform new updates and bug fixes with the field reconfigurability of the FPGA, without being subject to fixed hardware limitations. FPGAs, with their high performance and real-time data processing capability, offer an extremely broad portfolio that directs the use of in-vehicle artificial intelligence applications, which use cameras, radar, and LiDAR in automotive technologies, providing fully adaptive autonomous driving together with safety-critical functionality.

Broadcast

The prevalence of instant requests in today's communication systems has put this sector in search of hardware capable of advanced filtering and high-resolution transmission. Broadcast-targeted platforms and professional broadcast systems require applications that adapt quickly to changing requirements and have long product life cycles. The need for secure and high-performance hardware in many areas from television broadcasting to satellite broadcasting, from military surveillance to autonomous video processing brings FPGAs to the fore in this sector. FPGA provides the required remote configuration and long-life cycle form needed in space communication systems with its high processing capability and high bandwidth [35]. Many topics in this field such as broadcast vehicles, space observation vehicles, cinema, video processing, compression, archiving, transmission, and switching are different as manufacturers offer FPGAs as a solution for real-time video and audio processing applications required for all these videos and image-based systems. has gained an orientation. As a result of the studies in this field [36-38], the timing problems in satellite systems for digital video broadcasting, the need to minimize energy, and the transmission problems in high-resolution image systems, the fast resolution feature of FPGA have been solved.

Robotics

Robots and robotic systems reflect the current expression for the current industrial automation. Machine vision, communication, industrial control, artificial intelligence, machine learning, and human-machine interaction are important technological topics of industrial robot applications and robotics [39]. Application packages that manufacturers have developed specifically for the robotics industry are the simplest way to create robots with computational visualization, hardware acceleration, and machine vision systems. FPGAs offer a reconfigurable common embedded hardware and software platform to significantly reduce the time-to-market and cost of products developed for the robotics industry. This provides a modular approach in the field of robotics. Parallel processing of control parameters enables precise and dynamic industrial standards to be achieved in a scaled number of motion axes, various and multiple sensor support, and high-level sensor techniques can be applied [40-42].
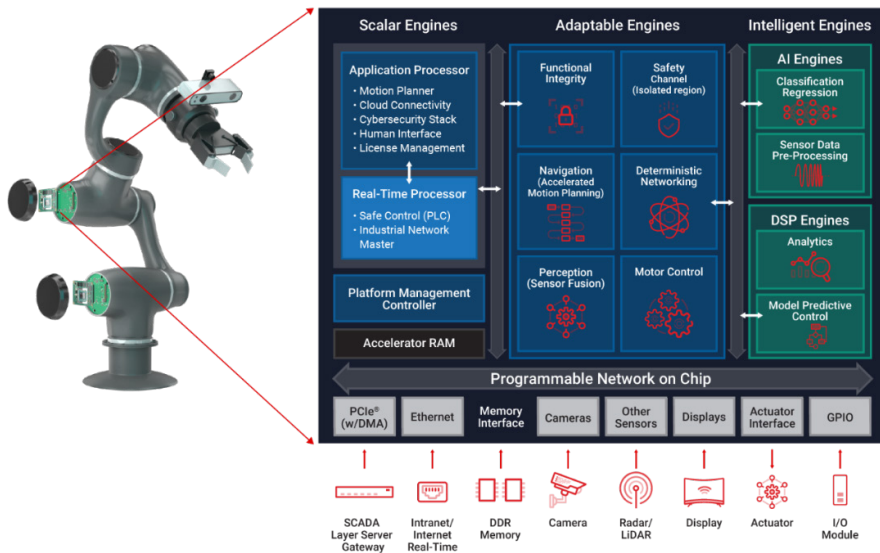
Figure 4. Robotics applications with FPGA [41]

With FPGA hardware, deterministic networking with low latency for system synchronization is achieved in motor control. Adaptive motor control accelerates the planning required for movement. In this way, with a dynamic control structure, and artificial intelligence supported, predictive maintenance analytics can be applied in real-time. Studies on motor control applications with FPGA hardware [43-45] have been carried out for high-loop operations such as direct torque control, sensorless control, and fuzzy logic-based control. The high-speed advantage of the FPGA has been observed to maximize the computational and feedback loops while minimizing torque ripple.

Data Center

With the increase in data size and operations with technological devices, the importance of cloud computing and cloud data distributions is also increasing. They can provide the required performance and capacity by being equipped with FPGAs for high-bandwidth, low-latency servers, fast networking, and fast storage applications. The flexibility of hardware reconfiguration, high throughput, and low latency have led developers to FPGAs for areas such as network acceleration, financial technology, and data processing applications. HPC applications, imaging, and video centers in general benefit from FPGA technologies due to their high throughput, low power consumption efficiency, and workload adaptability advantages. Researchers have conducted studies in data centers based on fixed-point computing performance to increase inter-node error reporting performance, processing efficiency, and acceleration [46, 47]. Their results show that FPGA and GPU perform at the same level, but FPGA achieves this performance with more than 50 times lower latency.

Medical

In addition to being a targeted design platform for industry, science, and medical (ISM), FPGAs provide a high level of flexibility, fast time-to-market, and non-repeating engineering costs (NRE) for industrial applications. In medicine, their use is common for medical imaging applications. In operations performed on images, their ability to produce fast results with their parallel processing capabilities ensures that diagnosis and diagnosis processes are accelerated. Scientific simulations, machine learning, and FPGA performance of arrays can meet the need in this field. They are similarly used in industrial applications to meet string processing, display, and I/O interface requirements.
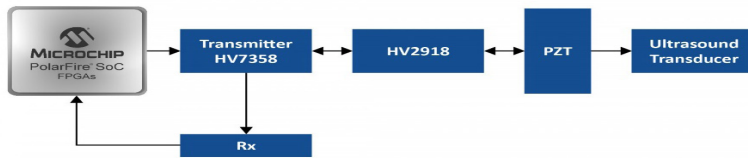
Figure 5. Medical ultrasound application with FPGA [12]

In general, studies with FPGA in the medical field [48-50] have focused on medical diagnosis identification with real-time data analysis, image processing and analysis on embedded systems, and ultrasound applications. It is stated that real-time monitoring is an important factor in providing detailed analysis during medical imaging.

Video, Image Processing, and Security

Many security applications, including access control, surveillance and security systems, and video and image processing systems need solutions that meet increasing mobility. In this context, FPGAs provide high performance versus high reconfigurability, low NRE, and low power consumption for a variety of targeted video and display applications. Today, area scanning cameras with resolutions that can exceed 20MP, high frame rate and 3D detection technologies require high bandwidth, multi-interface support, and low energy consumption with the inclusion of artificial intelligence in the application. Technology companies are trying to provide the industry with high-capacity, industry-specific FPGA hardware that can meet all needs on a single system to meet this need. However, researchers play a key role in determining the need by presenting various studies on this equipment. Recently, holographic video processing and real-time object identification have come to the fore among studies [48, 51, 52].
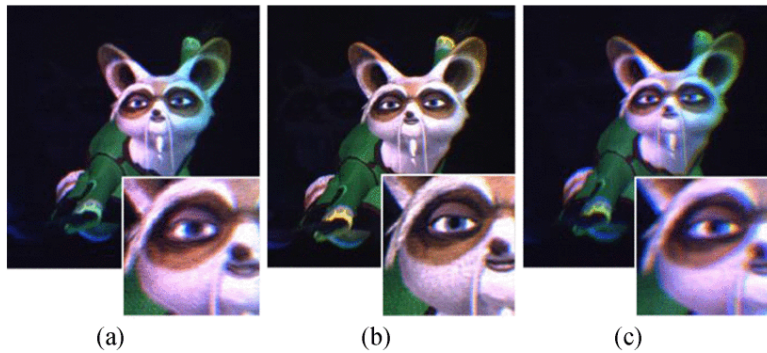
Figure 6. FPGA reconstructed holographic image (a) 0m, (b) 0.4m, and (c) 0.6m [52]

Communication

It can produce end-to-end solutions for reprogrammable network packages, serial panels, and more on FPGA hardware. They are widely used in the communications industry to increase network capacity while reducing delays, lower power consumption while increasing coverage, and improve connection quality, especially when it comes to data transmission.

Tech leaders such as Microsoft, Google, and Amazon have started using configurable FPGAs in their server systems to accelerate and add flexibility to various workloads. However, FPGAs need skilled human resources for both hardware and software design. For this reason, manufacturers and developers have started to work on customized IP packages to overcome the human resource barrier. IP packets are defined as libraries produced specifically for the application. With this development, companies have turned to the use of FPGAs in cloud services, network structures, and server systems. According to experts, the use of industrial applications built on this hardware is expected to increase further over the years [53].

CONCLUSIONS

The purpose of this article is to provide information about the usage areas and advantages of FPGAs in industrial applications. It has become important to use an efficient design architecture as the complexity of control methods in the industry increases depending on the devices used and the need. FPGA hardware is based on system-level design and system-level description language. This hardware allows rapid prototyping of applications. Verification of the design for the target technology during the application development phase provides the developer with a reliable structure for the application. With verification, a smooth development step is achieved at once. The availability of validated packages (IP block) for designs provides great advantages during the development phase. In another aspect, it allows the integration of digital controllers into complex systems as models.

According to the studies, it can be observed that FPGA hardware offers high performance compared to traditional processors when an application requires rapid processing of large amounts of data. Although the use of traditional processors is

still a flexible option in every field, the demanding amounts of data and increasing precision in today's applications make the use of FPGAs in almost every field a necessity.

FPGA technology can offer customizable and high-performance solutions with low energy consumption and less carbon footprint compared to other technologies that cannot provide an optimum solution for modern information processing techniques. In addition, it helps to bring reliable products to the market without the need for designers to take extra action for the functional safety of applications. Among the advantages that are considered common in applications, longevity, high bandwidth, functional security, flexibility and reconfigurability, high transaction performance and parallel processing capability come to the fore. Therefore, FPGAs are a viable choice for designers to meet the rapidly increasing and changing needs in the industrial applications of the new century.

The evaluation of the difficulties encountered in the past years by today's manufacturers and the solutions developed for these problems have produced positive effects such as cost reduction, ease of installation, and ease of integration. This has started to make FPGAs a pioneering solution for industrial applications. Intel and AMD Xilinx companies have been working on special applications of industrial solutions, mostly by providing improvements on FPGAs in recent years. They offer high-density roadmaps especially in industrial communications, industrial motor control, machine vision, video and broadcast, robotics, artificial intelligence, edge computing, and semiconductors.

Enhanced performance, low cost, faster time to market (QTTM), flexible and reliable structure, and low maintenance requirements make FPGAs advantageous in different industries compared to other processor types. In terms of production, they are classified into low, medium, and high segments. Low-segment hardware is used for systems that require intensive logic groups and complexity with low on-chip power consumption. Mid-range hardware is the ideal solution for systems that require moderate complexity and balance the consumption per hardware, without sacrificing both factors while maintaining a performance and cost ratio. High-segment hardware, on the other hand, offers solutions with its functionality and performance in applications where cost is not so important, consumes relatively more power, but requires high logic density and high complexity.

In the years to come, the complexity of the control systems and technologies used will continue to increase the amount of data that emerges. The tasks set for the applications will not be limited to editing and adaptive online control due to errors will need to be provided. In this context, FPGAs will continue to be the subject of research due to their dynamic configuration architecture and increasing performance need. By looking at the studies on this hardware and the roadmaps of the manufacturers, it can be predicted that the design difficulty and the need for special expertise for FPGAs will be eliminated in the future.

REFERENCES

[1] R. Skhiri, V. Fresse, J. P. Jamont, B. Suffran, and J. Malek, "From FPGA to support cloud to cloud of FPGA: State of the art," *International Journal of Reconfigurable Computing,* vol. 2019, 2019.

[2] J. Fowers, G. Brown, P. Cooke, and G. Stitt, "A performance and energy comparison of FPGAs, GPUs, and multicores for sliding-window applications," in *Proceedings of the ACM/SIGDA international symposium on Field Programmable Gate Arrays*, 2012, pp. 47-56.

[3] J. J. Rodríguez-Andina, M. D. Valdes-Pena, and M. J. Moure, "Advanced features and industrial applications of FPGAs—A review," *IEEE Transactions on Industrial Informatics,* vol. 11, no. 4, pp. 853-864, 2015.

[4] J.-X. Zhang, F.-S. Chen, J. Cheng, and X.-J. Bian, "Study of implementing the FEC on SoC FPGA for digital quench detector," *Radiation Detection Technology and Methods,* vol. 2, no. 1, pp. 1-8, 2018.

[5] I. Firmansyah, Y. N. Wijayanto, and Y. Yamaguchi, "2D stencil computation on Cyclone V SoC FPGA using OpenCL," in *2018 International Conference on Radar, Antenna, Microwave, Electronics, and Telecommunications (ICRAMET)*, 2018: IEEE, pp. 121-124.

[6] M. Elnawawy, A. Farhan, A. Al Nabulsi, A.-R. Al-Ali, and A. Sagahyroon, "Role of FPGA in internet of things applications," in *2019 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*, 2019: IEEE, pp. 1-6.

[7] Y. Ma, P. Duan, P. He, F. Zhang, and H. Chen, "FPGA implementation of extended Kalman filter for SOC estimation of lithium-ion battery in electric vehicle," *Asian Journal of Control,* vol. 21, no. 4, pp. 2126-2136, 2019.

[8] S. Yang *et al.*, "Research on Heterogeneous Cloud Test Platform Based on Elastic Scaling Mechanism," in *2019 IEEE 19th International Conference on Software Quality, Reliability and Security Companion (QRS-C)*, 2019: IEEE, pp. 493-498.

[9] S. Bouhoun, R. Sadoun, and M. Adnane, "OpenCL implementation of a SLAM system on an SoC-FPGA," *Journal of Systems Architecture,* vol. 111, p. 101825, 2020.

[10] A. Kojima, "Autonomous Driving System implemented on Robot Car using SoC FPGA," in *2021 International Conference on Field-Programmable Technology (ICFPT)*, 2021: IEEE, pp. 1-4.

[11] G. View, "Field Programmable Gate Array Market Size, Share & Trends Analysis Report By Technology (SRAM, Antifuse, Flash), By Application (Military & Aerospace, Telecom), By Region, And Segment Forecasts," https://www.grandviewresearch.com/industry-analysis/, 978-1-68038-133-7, 2020.

[12] Microchip. "FPGAs for Industrial Applications." Microchip. (accessed November 2022).

[13] *XQ Defense-Grade Versal Portfolio*, https://www.xilinx.com/content/dam/xilinx/publications/product-briefs/xilinx-xq-versal-product-brief.pdf, 2018.

[14] Intel. "Industrial Solutions Playbook for Intel® FPGAs." Intel. (accessed November 2022).

[15] A. Xilinx. "What is an FPGA?" Xilinx. https://www.xilinx.com/products/silicon-devices/fpga/what-is-an-fpga.html (accessed November 2022).

[16] E. Monmasson, L. Idkhajine, M. N. Cirstea, I. Bahri, A. Tisan, and M. W. Naouar, "FPGAs in industrial control applications," *IEEE Transactions on Industrial informatics,* vol. 7, no. 2, pp. 224-243, 2011.

[17] А. Поляков, Языки VHDL и VERILOG в проектировании цифровой аппаратуры. Litres, 2022.

[18] Y. Lin. (2013) FPGA devices for solving challenges in industrial applications. *EngineerIT.*

[19] H. Amano, *Principles and structures of FPGAs*. Springer, 2018.

[20] A. Boutros and V. Betz, "FPGA architecture: Principles and progression," *IEEE Circuits and Systems Magazine,* vol. 21, no. 2, pp. 4-29, 2021.

[21] P. Lei, J. Liang, Z. Guan, J. Wang, and T. Zheng, "Acceleration of FPGA based convolutional neural network for human activity classification using millimeter-wave radar," *IEEE Access,* vol. 7, pp. 88917-88926, 2019.

[22] S. Janković *et al.*, "High-capacity FPGA router for satellite backbone network," *IEEE Transactions on Aerospace and Electronic Systems,* vol. 56, no. 4, pp. 2616-2627, 2019.

[23] Z.-L. Tang, S.-M. Li, and L.-J. Yu, "Implementation of deep learning-based automatic modulation classifier on FPGA SDR platform," *Electronics,* vol. 7, no. 7, p. 122, 2018.

[24] A. Toubal, B. Bengherbia, M. O. Zmirli, and A. Guessoum, "FPGA implementation of a wireless sensor node with built-in security coprocessors for secured key exchange and data transfer," *Measurement,* vol. 153, p. 107429, 2020.

[25] K. Cosmas and A. Kenichi, "Utilization of FPGA for onboard inference of landmark localization in CNN-Based spacecraft pose estimation," *Aerospace,* vol. 7, no. 11, p. 159, 2020.

[26] R. Al-Alwani, M. Al-Helo, and S. Alwash, "Enhancement In FPGA Using Moving Target Defence From Security Threats," *Solid State Technology,* pp. 3083-3093, 2020.

[27] S. Liu, D. Li, W. Shen, Z. Wang, G. Yang, and X. Song, "Application Research of Formal Verification in Aerospace FPGA," in *2021 IEEE 21st International Conference on Software Quality, Reliability and Security Companion (QRS-C)*, 2021: IEEE, pp. 797-805.

[28] S. Jung and J. P. Choi, "Predicting system failure rates of SRAM-based FPGA on-board processors in space radiation environments," *Reliability Engineering & System Safety,* vol. 183, pp. 374-386, 2019.

[29] P. Mohankumar, J. Ajayan, R. Yasodharan, P. Devendran, and R. Sambasivam, "A review of micromachined sensors for automotive applications," *Measurement,* vol. 140, pp. 305-322, 2019.

[30] J. H. Oh *et al.*, "An FPGA-based electronic control unit for automotive systems," in *2019 IEEE International Conference on Consumer Electronics (ICCE)*, 2019: IEEE, pp. 1-2.

[31] B. Jiang, N. Sharma, Y. Liu, C. Li, and X. Huang, "Real-Time FPGA/CPU-Based Simulation of a Full-Electric Vehicle Integrated with a High-Fidelity Electric Drive Model," *Energies,* vol. 15, no. 5, p. 1824, 2022.

[32] I. Vido, I. Škorić, D. Mitrović, M. Milošević, and M. Herceg, "Automotive Vision Grabber: FPGA design, cameras and data transfer over PCIe," in *2019 Zooming Innovation in Consumer Technologies Conference (ZINC)*, 2019: IEEE, pp. 103-108.

[33] K. Cho, J. Kim, D. Y. Choi, Y. H. Yoon, J. H. Oh, and S. E. Lee, "An FPGA-based ECU for remote reconfiguration in automotive systems," *Micromachines,* vol. 12, no. 11, p. 1309, 2021.

[34] H. Bai, H. Luo, C. Liu, D. Paire, and F. Gao, "Real-time modeling and simulation of electric vehicle battery charger on FPGA," in *2019 IEEE 28th International Symposium on Industrial Electronics (ISIE)*, 2019: IEEE, pp. 1536-1541.

[35] N. Singh, S. Chetan, K. Sourabh, and J. Manikandan, "FPGA Implementation of 32APSK Demodulator for Digital Video Broadcasting DVB-S2," in *2021 IEEE Mysore Sub Section International Conference (MysuruCon)*, 2021: IEEE, pp. 638-642.

[36] P. V. dos Santos, J. C. Alves, and J. C. Ferreira, "An FPGA array for cellular genetic algorithms: Application to the minimum energy broadcast problem," *Microprocessors and Microsystems,* vol. 58, pp. 1-12, 2018.

[37] L. Geng *et al.*, "P-32: 4K-UHD Broadcast Monitor Display System Based on FPGA," in *SID Symposium Digest of Technical Papers*, 2021, vol. 52, no. 1: Wiley Online Library, pp. 1182-1185.

[38] L. Guo *et al.*, "Analysis and optimization of the implicit broadcasts in FPGA HLS to improve maximum frequency," in *2020 57th ACM/IEEE Design Automation Conference (DAC)*, 2020: IEEE, pp. 1-6.

[39] Z. Wan *et al.*, "A survey of FPGA-based robotic computing," *IEEE Circuits and Systems Magazine,* vol. 21, no. 2, pp. 48-74, 2021.

[40] A. Podlubne and D. Göhringer, "FPGA-ROS: Methodology to augment the robot operating system with FPGA designs," in *2019 International Conference on ReConFigurable Computing and FPGAs (ReConFig)*, 2019: IEEE, pp. 1-5.

[41] A. Xilinx. "Robotics: Modular Platforms for Multi-Axis, Vision-Guided, and Safety-Enabled Intelligent Robotics." Xilinx. https://www.xilinx.com/applications/industrial/robotics.html (accessed November 2022).

[42] J. P. Queralta *et al.*, "FPGA-based architecture for a low-cost 3D lidar design and implementation from multiple rotating 2D lidars with ROS," in *2019 IEEE SENSORS*, 2019: IEEE, pp. 1-4.

[43] H. M. A. Abbas, R. F. Chisab, and M. J. Mnati, "Monitoring and controlling the speed and direction of a DC motor through FPGA and comparison of FPGA for speed and performance optimization," *International Journal of Electrical & Computer Engineering (2088-8708),* vol. 11, no. 5, 2021.

[44] A. Aib, D. E. Khodja, S. Chakroune, and L. Benyettou, "On-Line FPGA Hardware in the Loop Validation of Based Fuzzy-STPWM Induction Motor Control," *Journal homepage: http://iieta.org/journals/ama_b,* vol. 65, no. 1-4, pp. 17-26, 2021.

[45] S. Krim, S. Gdaim, A. Mtibaa, and M. Faouzi Mimouni, "FPGA-based real-time implementation of a direct torque control with second-order sliding mode control and input–output feedback linearisation for an induction motor drive," *IET Electric Power Applications,* vol. 14, no. 3, pp. 480-491, 2020.

[46] Y. Cai *et al.*, "FPGA investigation on error-flare performance of a concatenated staircase and Hamming FEC code for 400G inter-data center interconnect," *Journal of Lightwave Technology,* vol. 37, no. 1, pp. 188-195, 2018.

[47] X. Yu *et al.*, "A data-center FPGA acceleration platform for convolutional neural networks," in *2019 29th International Conference on Field Programmable Logic and Applications (FPL)*, 2019: IEEE, pp. 151-158.

[48] A. Sanaullah, C. Yang, Y. Alexeev, K. Yoshii, and M. C. Herbordt, "Real-time data analysis for medical diagnosis using FPGA-accelerated neural networks," *BMC bioinformatics,* vol. 19, no. 18, pp. 19-31, 2018.

[49] Y. Wang *et al.*, "Developing medical ultrasound beamforming application on GPU and FPGA using oneAPI," in *2021 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW)*, 2021: IEEE, pp. 360-370.

[50] T. Wu and Z. Yang, "Animal tumor medical image analysis based on image processing techniques and embedded system," *Microprocessors and Microsystems,* vol. 81, p. 103671, 2021.

[51] R. Solovyev, A. Kustov, D. Telpukhov, V. Rukhlov, and A. Kalinin, "Fixed-point convolutional neural network for real-time video processing in FPGA," in *2019 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus)*, 2019: IEEE, pp. 1605-1611.

[52] H. Kim *et al.*, "A single-chip FPGA holographic video processor," *IEEE Transactions on Industrial Electronics,* vol. 66, no. 3, pp. 2066-2073, 2018.

[53] K. Freund. (2018) An Update On Data Center FPGAs.

DEVELOPMENT OF ENERGY SAVING AND ECO FRIENDLY COMMERCIAL DISHWASHERS USING OZONE

Zafer KAHRAMAN [1], Murat HACI [2], Hakan Serhad SOYHAN [3]

INTRODUCTION

Ozone (O_3) has been expanding its usage areas in recent years with its strong oxidizing and disinfectant properties [1]. Ozone applications for disinfection of different product groups are encountered in various sectors such as food [2-9], wastewater treatment [10-15], textiles [16-19] and medical fields [20-22]. Summaries of various literature studies on the applications of ozone in different fields are given below.

Many parameters affect ozone applications. The main environmental factors affecting the ozonation process applied as a gas are temperature, pressure, relative humidity, etc. The effectiveness of aqueous ozone applications varies across many factors (eg pH, conductivity and organic matter composition, etc.). Under ambient conditions, material type and properties (adsorbent or non-adsorbent), microorganism structure, sample contamination method (droplets, wet, dried), ozone production method, ozone concentration, and exposure time stand out as other factors affecting the efficiency of ozonation. The affecting factors are given in Figure 1. under three main groups (ambient conditions, material/substrate, and operational conditions) [1].
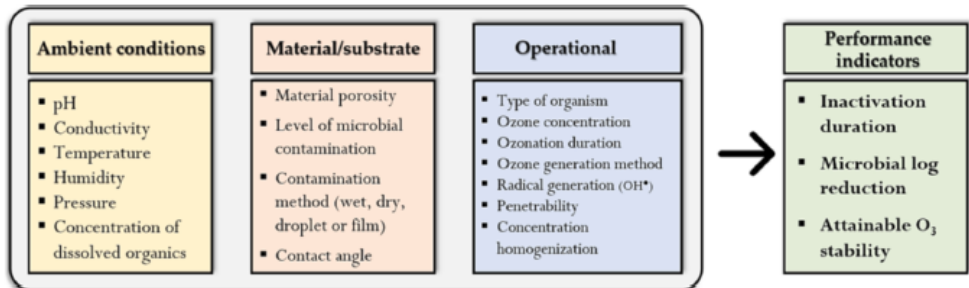
Figure 1. Classification of parameters affecting the ozonation process. [1]

In the ozonation process, negative results may be encountered instead of obtaining the desired effect if the application parameters are not determined correctly. In determining the best parameters for ozone applications, excessive and low ozone

dose applications should be avoided. The ozone concentration ranges (Occupational Safety and Health Administration) determined for different applications are given in Figure 2. The contact time and concentration of ozone stand out as important factors in the application process [1].
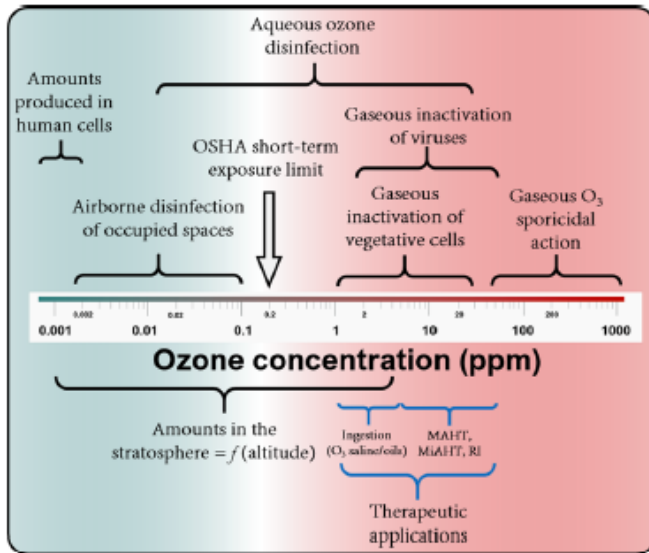
Figure 2. The exposure limits for various applications to naturally occurring ozone and commonly applied ozone concentrations. While ozone application generally has tiny concentrations of 1-3 mg/L for drinking water disinfection, Up to 10 mg/L can be applied for wastewater disinfection (although its organic content varies). Doses higher than those specified for industrial wastes are also applied [1].

Emmanuel his working group evaluated the importance of effective determination of ozone dosage requirements (concentration x exposure time) in order to maximize efficiency with a suitable technique in the control of ozone applications [1].

Mustapha and his work group studied the effects of low concentration, multi-mode frequency irradiation of aqueous ozone and their combination on cherry tomatoes' microbial safety and nutritional quality. In this study, individual washing with aqueous ozone, single-mode frequency irradiation reduced (<1 log CFU/g) in microorganisms that cause deterioration. Higher microbial reduction (1.3 to 2.6 1 log CFU/g) was obtained with dual-mode frequency irradiation (DMFI). The highest level (>3 log CFU/g) microbial reduction was achieved with the combined system (20-40 kHz and aqueous ozone) application. The representation of the aqueous ozone generator system, the multi-mode frequency irradiation system, and the sonozonation system is given in Figure 3 [2].
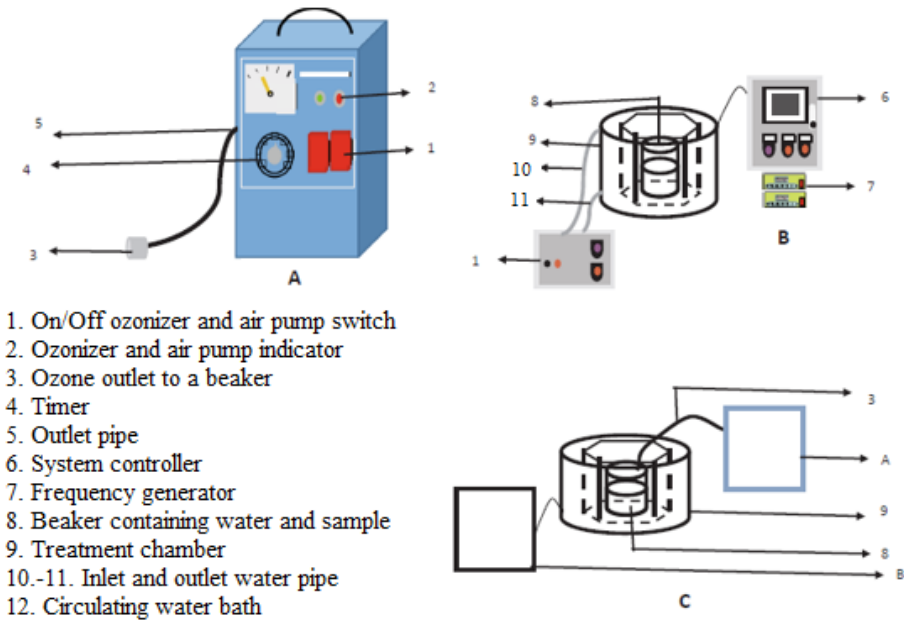
1. On/Off ozonizer and air pump switch
2. Ozonizer and air pump indicator
3. Ozone outlet to a beaker
4. Timer
5. Outlet pipe
6. System controller
7. Frequency generator
8. Beaker containing water and sample
9. Treatment chamber
10.-11. Inlet and outlet water pipe
12. Circulating water bath

Figure 3. A) Aqueous ozone generator, B) Multi-mode frequency irradiation system, C) Sonozonation system [2].

Zambre and his working group researched the shelf life of tomatoes with ozone applications. The effects of ozone application on tomatoes' storage and ripening processes were evaluated. The color changes of tomatoes were investigated (from green to red). They stated that ozone applications delayed the development of tomatoes (the red color), rot and increased the shelf life of tomatoes by 12 days. The illustration of the experimental system developed for ozone application is given in Figure 4 [3].
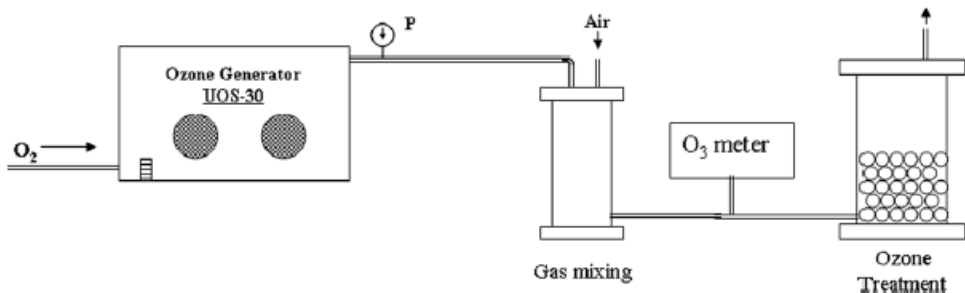
Figure 4. The experimental system for ozone application. To obtain the desired ozone concentration utilizing an ozone generator, ozone was mixed in the gas chamber, an ozone meter was used for measurement. The gas inlet containing the desired ozone concentration was continuously passed through a tomato stack in the ozone treatment chamber for 10 min. In addition, tomatoes' redness and rotting values were obtained during storage [3].

Joseph and his working group evaluated the ozonation process of wastewater containing laundry detergent and made suggestions for next studies. The most effective advanced oxidation processes (AOPs), ozonation stands out with its effectiveness in treating and recreating laundry detergent wastewater. Ozonation processes are generally evaluated under two separate headings as non-catalytic ozonation and catalytic ozonation. The representation of the ozonation is given in Figure 5 [4].
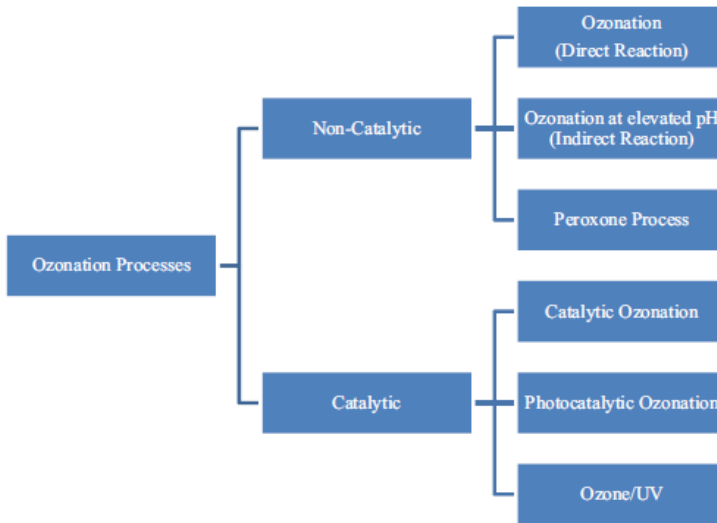
Figure 5. Schematic of various ozonation processes. [4].

Organic compounds (such as lignin, cellulose, phenol) and other toxic compounds wastewater cause an increase in chemical oxygen demand (COD), color, and properties related to suspended solids. In addition, the high biological material and color content in paper and pulp wastewater can lead to the formation of sludge and scum and lower visual quality in water bodies. Ozone-based techniques are contributing as an alternative among AOPs to remove persistent organic compounds. The most appropriate ozone dose amount (0.3 mg/ml) has been reported for COD and color removal of wastewater. The illustration of the catalytic ozonation system (for pulp and paper industry wastewater treatment) is given in Figure 6 [13].
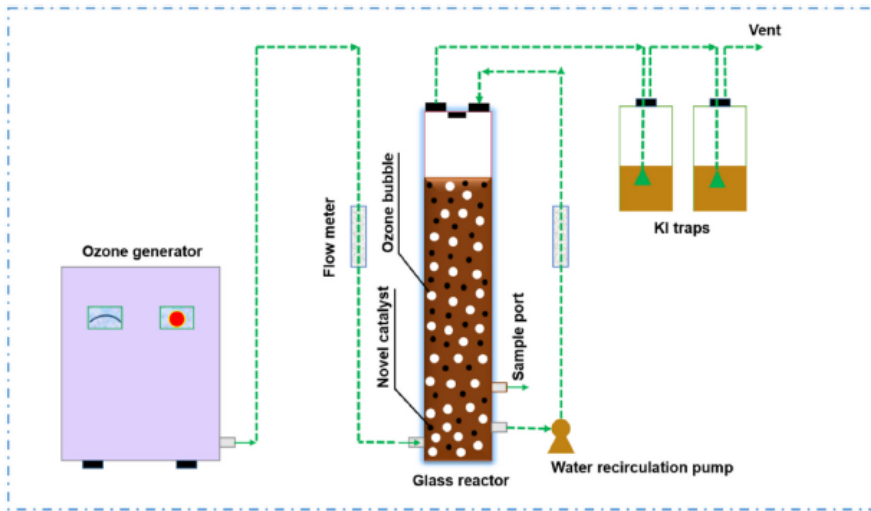
Figure 6. The catalytic ozonation system [13].

The dirty image of the wastewater sample and the final image formed after the ozonation process is given in Figure 7 [13].
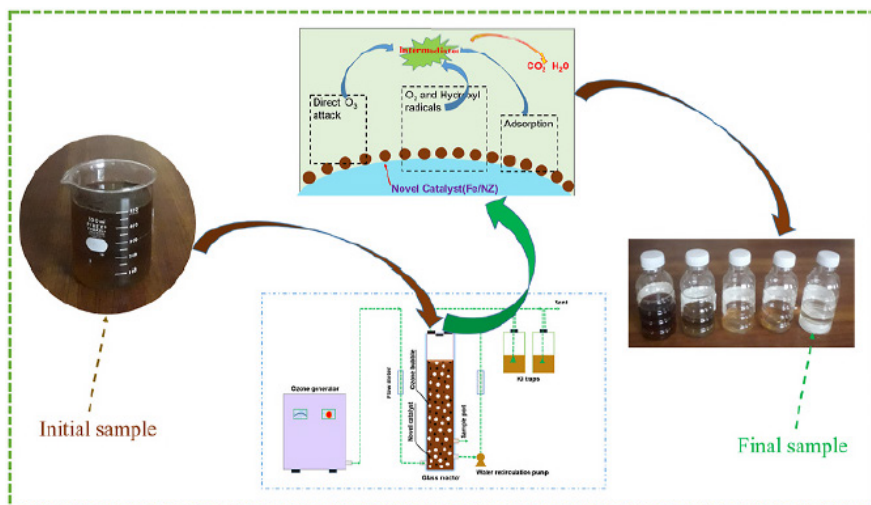
Figure 7. Image of initial sample and final image after ozonation [13].

Shajeelammal and his work group investigated ozonation (O_3) and modified low-frequency ultrasound (U.S.) cavitation as well as the integrated reactor for possible replacement of the stage-1 plant that produces a large amount of sludge in textile production. According to the company's needs, activities were carried out to eliminate at least 90% of the chemical oxygen demand (COD) without forming sludge, with the COD level below 150 mg L^{-1} [16].

The change in COD of textile wastes as a function of ozonation time is given in Figure 8 [16].
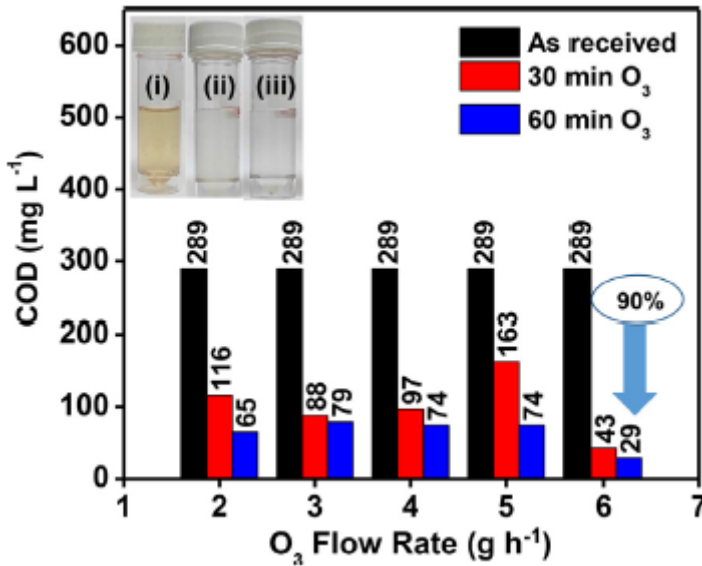
Figure 8. Variation in the COD values, obtained for various ozonation (times, flow rates etc.). The inset shows photographs of textile effluent in the as-received condition (i) and after the ozonation time of 30 (ii) and 60 min (iii) for an optimum O_3 flow rate of 6 g h^{-1} [16].

The coronavirus pandemic has caused adverse events worldwide. Bayarri and his team reported that technologies based on ozone (has a strong oxidizing property), eliminated the activity of this virus in various structures (aerosols and fomites). The illustration of the disruption of activation mechanisms for viruses (enveloped and non-enveloped) is given in Figure 9 [20].
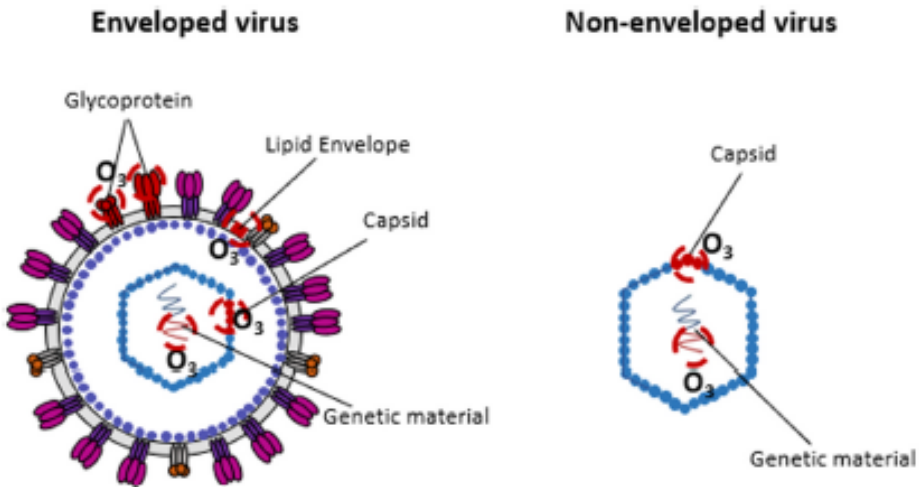
Figure 9. Inactivation mechanisms of viruses (enveloped and non-enveloped) by ozone [20].

METHODS

Commercial dishwashers are used in commercial establishments (restaurants, shopping malls, hotels, dormitories, etc.) to wash large amounts of dirty products (plates, glasses, etc.) in a short time (maximum 3 minutes according to the washing program).

Commercial dishwashers are used in various models (glass washing, undercounter, guillotine type, conveyor, etc.) and capacities (500, 1080, 1500, 3000 plates/hour etc.). Commercial enterprises choose the model and capacity of dishwashers according to the conditions. In this study, an innovative commercial dishwasher prototype using ozone was developed based on the hood type commercial dishwasher.

Consumer demands are increasing for using environmentally friendly water and detergent/rinse aid in washing processes in commercial dishwashers and reducing energy consumption. Depending on the amount of chemicals (detergents and rinse aid) used in the existing washing units, it may pose great risks for the environment after the washing process.

The innovative prototype were carried out effectively by the expert staff of our R&D Center. Initially, the original design of the innovative commercial dishwasher prototype using ozone was developed. Stainless steel (AISI 316) is used in the main body of the prototype's washing chamber and critical equipment (washing tank, washing arms, etc.) in order to provide high resistance in the process of working with ozone. While polymer-based materials are used in the washing arms of standard commercial dishwashers, only AISI 316 quality material is used in the innovative prototype.

The performance of an environmentally friendly innovative commercial dishwasher using ozone in different usage conditions was investigated. Various reviews and evaluations on the innovative prototype are given below;

- Working conditions of prototype commercial dishwasher (washing time and washing efficiency etc.),
- Working in different ozone application conditions (ozone application dose, ozone application temperature, time, etc.) in the prototype commercial dishwasher,
- Working with occupational safety of prototype commercial dishwasher,
- Various ozone application times,
- Effects of different washing programs,
- Compatibility of the prototype commercial dishwasher with ozone equipment,
- Mechanical balance and strength of the prototype commercial dishwasher,
- Various material qualities used in the prototype commercial dishwasher,
- Test and analysis evaluations of various washed products (plates, glasses etc.)

- The hygiene conditions (total bacteria, e.coli etc.) of the products washed using ozone in the innovative system.

An ozone generator with a 1000 mg/hour capacity with a pump that produces ozone by arc method was used. As the washing programs, the plate is selected in the slightly soiled position and the washing time is 40 seconds; it is set as 2 seconds filtering time and 10 seconds rinsing time. Standard washing system test conditions are given below;

- The commercial dishwasher prepared under standard conditions is equipped with detergent and rinse aid dosing pumps.

- The rinse aid pump dosage adjustment is 1 gram/liter.

- The washing water temperature of the commercial dishwasher prototype is set to 55°C.

- The rinse water temperature of the commercial dishwasher prototype is set to 85°C.

- Detergent pump dosage adjustment is made as 3 grams/liter.

Innovative eco friendly commercial dishwasher system using ozone is given in Figure 10.

Figure 10. Innovative eco friendly commercial dishwasher system using ozone (front view).

CONCLUSION

Ozone systems are used in various areas, food preservation, hygiene (cold air warehouse, water filling facilities, etc.), disinfection (water treatment, laundry, etc.) and wastewater treatment (turbidity removal, etc.) due to its many superior properties (its oxidation power being more effective than many disinfectants). In

the specially designed commercial dishwasher prototype, the superior properties of ozone contributed to the reduction of the use of detergent and rinse aid. The effective reaction temperatures of ozone are lower than the current washing temperature (~55-60°C), and gains have been achieved in reducing energy consumption values. In addition, different bacteria (e.coli, etc.) were removed at the desired level in various kitchen products (plates, glasses, etc.) tested by contaminating the dishwasher prototype using ozone.

An innovative prototype was successfully obtained based on the R&D systematic, by effectively evaluating the scientific and technical outputs for eliminating risks/uncertainties on the commercial dishwasher prototype using ozone.

ACKNOWLEDGEMENTS

REFERENCES

[1] Emmanuel I. E., Macfarlane, A., Cusack, M., Burns, A., Okolie, J.A., Mackay, W., Rateb, M., Yaseen. M. Ozone application in different industries: A review of recent developments. Chemical Engineering Journal, **2023**, 454, 140188.

[2] Mustapha AT, Zhou C, Wahia H, Amanor-Atiemoh R, Out P, Qudus A, Fakayode OA, Ma H. Sonozonation: Enhancing the antimicrobial efficiency of aqueous ozone washing techniques on cherry tomato. Ultrasonics – Sonochemistry, **2020**, 64,105059.

[3] Zambre, S.S., Venkatesh, K.V., Shah, N.G. Tomato redness for assessing ozone treatment to extend the shelf life. Journal of Food Engineering, **2010**, 96, 463–468.

[4] Joseph, C.G., Farm, Y.Y., Taufiq-Yap, Y.H., Pang, C.K., Nga, J.L.H., Puma, G.L. Ozonation treatment processes for the remediation of detergent wastewater: A comprehensive review. Journal of Environmental Chemical Engineering, **2021**, 9, 106099.

[5] Niveditha, A., Pandiselvam, R., Prasath, V.A., Singh, S.K., Gul, K., Kothakota. A. Application of cold plasma and ozone technology for decontamination of Escherichia coli in foods- a review. Food Control, **2021**, 130, 108338.

[6] Paulikien, S., Venslauskas, K., Raila, A., Zvirdauskien, R., Naujokien, V. The influence of ozone technology on reduction of carrot loss and environmental IMPACT. Journal of Cleaner Production, **2020**, 244, 118734.

[7] Pandiselvam, R., Singh, A., Agriopoulou, S., Sachadyn-Krol, M., Aslam, R., Lima, C.M.G., Khanashyam, A.C., Kothakota, A., Atakan, O., Kumar, M., Mathanghi, S.K., Khaneghah, A.M. A comprehensive review of impacts of ozone treatment on textural properties in different food products. Trends in Food Science & Technology, **2022**, 127, 74–86.

[8] Castanha, N., Junior, M.D.M., Augusto, P.E.D. Potato starch modification using the ozone technology. Food Hydrocolloids, **2017**, 66, 343-356.

[9] Kaur, K., Pandiselvam, R., Kothakota, A., Ishwarya, S.P., Zalpouri, R., Mahanti, N.K. Impact of ozone treatment on food polyphenols – A comprehensive review. Food Control, **2022**, 142, 109207.

[10] Wang, B., Liu, Y., Zhang, H., Shi, W., Xiong, M., Gao, C., Cui, M. Hydrodynamic cavitation and its application in water treatment combined with ozonation: A review. Journal of Industrial and Engineering Chemistry, **2022**, 114, 33–51.

[11] Zhang, H., Zhang, Y., Qiao, T., Hu, S., Liu, J., Zhu, R., Yang, K., Li, S., Zhang, L. Study on ultrasonic enhanced ozone oxidation of cyanide-containing wastewater. Separation and Purification Technology, **2022**, 303, 122258.

[12] Gijn, K.V., Zhao, Y., Balasubramaniam, A., Wilt, H.A., Carlucci, L., Langenhoff, A.A.M., Rijnaarts, H.H.M. The effect of organic matter fractions on micropollutant ozonation in wastewater effluents. Water Research, **2022**, 222, 118933.

[13] Munir, H.M.S., Feroze, N., Ramzan, N., Sagir, M., Babar, M., Tahir, M.S., Shamshad, J., Mubashir, M., Khoo, K.S. Fe-zeolite catalyst for ozonation of pulp and paper wastewater for sustainable water resources. Chemosphere, **2022**, 297, 134031.

[14] Lei, R., Shi, Y., Wang, X., Tao, X., Zhai, H., Chen, X. Water purification system based on self-powered ozone production. Nano Energy, **2021**, 88, 106230.

[15] Clem, V., Mendonça, H.V. Ozone reactor combined with ultrafiltration membrane: A new tertiary wastewater treatment system for reuse purpose. Journal of Environmental Management, **2022**, 315, 115166.

[16] Shajeelammal, J., Mohammed, S., Prathish, K.P., Jeeva, A., Asok, A., Shukla, S. Treatment of real time textile effluent containing azo reactive dyes via ozonation, modified pulsed low frequency ultrasound cavitation, and integrated reactor. Journal of Hazardous Materials Advances, **2022**, 7, 100098.

[17] Zhang, J., Mirza, N.R., Huang, Z., Du, E., Peng, M., Shan, G., Wang, Y., Pan, Z., Ling, L., Xie, Z. Evaluation of direct contact membrane distillation coupled with fractionation and ozonation for the treatment of textile effluent. Journal of Water Process Engineering, **2021**, 40, 101789.

[18] Hu, E., Shang, S., Tao, X., Jiang, S., Chiu, K. Regeneration and reuse of highly polluting textile dyeing effluents through catalytic ozonation with carbon aerogel catalysts. Journal of Cleaner Production, **2016**, 137, 1055-1065.

[19] Pazdzior, K., Wrebiak, J., Klepacz-Smołka, A., Gmurek, M., Bilinska, L., Kos, L., Sojka-Ledakowicz, J., Ledakowicz, S. Influence of ozonation and biodegradation on toxicity of industrial textile wastewater. Journal of Environmental Management, **2017**, 195, 166-173.

[20] Bayarri, B., Cruz-Alcalde, A., Lopez-Vinent, N., Mico, M.M., Sans, C. Can ozone inactivate SARS-CoV-2? A review of mechanisms and performance on viruses. Journal of Hazardous Materials, **2021**, 415, 125658.

[21] Ling, C., Liu, T., Li, R., Wang, S., Zhang, J., Li, S., Xia, L., Cui, M., Wang, P. A novel DCM-based NIR fluorescent probe for detecting ozone and its bioimaging in live cells. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, **2021**, 248, 119192.

[22] Franke, G., Knobling, B., Brill, F.H., Becker, B., Klupp, E.M., Campos, C.B., Pfefferle, S., Lutgehetmann, M., Knobloch, J.K. An automated room disinfection system using ozone is highly active against surrogates for SARS-CoV-2. Journal of Hospital Infection, **2021**, 112, 108-113.

EFFECTS OF DIETHYL ETHER ADDITION TO ISOPROPANOL-ISOBUTANOL-ETHANOL (IBE) AND DIESEL FUEL BLENDS ON PERFORMANCE AND EMISSIONS UNDER VARYING LOADS And SPEEDS

Nurullah GULTEKIN[1], Halil Erdi GULCAN[2], Murat CINIVIZ[3]

INTRODUCTION

Compression-ignition engines have better torque output and fuel efficiency because they can operate at higher compression ratios than spark-ignition engines. Therefore, a diesel engine is preferred for heavy construction machinery, sea transportation, public transportation vehicles, and agricultural machinery [1]. Moreover, HC, CO, and CO_2 pollutants are lower than spark ignition engine pollutants. However, despite these advantages, diesel engines also have negative sides. Diesel engines generate more NO_x and soot pollutants than spark ignition engines [2]. Pollutants from diesel combustion have been identified by the IARC as evidence that may increase people's risk of lung cancer. [3]. In addition, smoke pollutants from diesel combustion can cause cardiovascular disorders [4]. In addition, pollutants from diesel combustion can cause many harms, involving acid rain, and damage to the ozone layer, apart from human health [5, 6]. Improving diesel output emissions and decreasing reliance on fossil-based fuels is an important issue.

Many countries are dependent on foreign sources for fuel. Foreign dependency is an important parameter affecting economic development. Production of alternative fuels in the cheapest way and their use in internal combustion engines have become necessary for developing countries [7]. Finding alternatives to petroleum-based fuels reduces foreign dependency and helps economic development [8]. Nowadays, alcoholic fuels are used in addition to diesel fuel in different volumetric ratios to improve pollutant emissions without a serious decrease in performance [9, 10].

UTILISE OF ETHANOL IN A CI ENGINE

Ethanol is a captivating potential fuel since it has unlimited renewable bio-based raw materials and it has oxygen content, therefore ensuring the ability to decrease smoke emissions in a CI engines [11]. Ethanol is considered a regenerable fuel as it may be produced from agricultural crops such as corn, sugarcane, beets, and waste biomass [12, 13]. Ethanol is suitable as a fuel admixtion for working in a CI engine

1   Nurullah Gültekin, Teaching Assistant, Karamanoğlu Mehmetbey University, 0000-0002-0139-1352
2   Halil Erdi Gulcan, Research Assistant, Selcuk University, 0000-0002-2328-5809
3   Murat Ciniviz, Professor Doctor, Selcuk University, 0000-0003-3512-6730

with a high compression ratio owing to its high-octane content. The addition of ethanol significantly reduces smoke emissions thanks to the oxygen it contains [14]. It can also be used as an admixtion to enhance properties like distillation and cold flow [15, 16]. In the current literature, there are many studies on ethanol-diesel mixtures. In general, the adjunct of ethanol to diesel was found to have a notable impact on smoke pollutants. Some studies on ethanol-diesel are summarized below. Huang et al. [17] analysed the effects of adding 10%, 20%, 25%, and 30% ethanol to diesel on parameters of performance and emissions. It has been stated that owing to the lower calorific point of ethanol compared to diesel, the thermal efficiency decreases, and the specific fuel consumption increases. Otherwise, it has been informed that there is a notable attenuation in smoke emissions with the addition of ethanol to diesel. The reason for this is the oxygen concentration of ethanol. It was stated that the desired reduction in HC, CO, and NO_x emissions could not be achieved. The study by Ahmed focused on the ethanol additive in a compression-ignition heavy-duty engine. According to the test results, particulate matter emissions decreased by over 40%. And also, CO and NO_x emissions showed a reduction of over 25% and 5%, respectively [18]. Xing-Cai et al. [19] analysed the impact of cetane improver adjunct to ethanol diesel blends on high-speed diesel engine emissions and performance. Ethanol was added to diesel at a rate of 15% by volume. Cetane improver was added at 0%, 0.2% and 0.4% ratios. Ethanol was added to diesel at a rate of 15% by volume. Cetane improver was added at 0%, 0.2% and 0.4% ratios. Engine performance results indicated that specific fuel consumption increased, and thermal efficiency enhanced. Emissions, on the other hand, have improved. Smoke and NO_x emissions have decreased significantly. Bilgin et al. [20] focused on the effect of ethanol adjunct to diesel on performance at varying compression ratios. Ethanol was added to diesel in three different volumetric ratios (2%, 4% and 6%). Compression ratios in the study vary 19:1, 21:1 and 23:1. The performance results showed that the fuel consumption and thermal efficiency of the engine were improved at 4% ethanol substitution at different compression ratios. Bang-Quan He et al. [21] analysed the impacts of 10% and 30% ethanol mixtures by volume on pollutant emissions in a four-cylinder, direct injection, 17.5:1 compression ratio diesel engine. The tests were practiced at a constant 1700 rpm. The results indicated that the adjunct of ethanol significantly reduced soot emissions. Zhang et al. [22] worked the solubility of ethanol in diesel and the effect of its mixtures on emissions. Experiments were conducted with fuels containing different mixtures in a direct injection CI engine. The ethanol ratios in the mixture were 10%, 20%, and 30% by volume. Zhang et al. beholden a significant advance in smoke, HC, and CO emissions with the adjunct of ethanol.

USE OF ISOPROPANOL IN DIESEL ENGINES

Isopropanol which oxygenated content is another potential alternative fuel that can also be generated from raw materials as ethanol. Isopropanol provides better resolution in diesel blends due to its longer chain structure compared to ethanol [23]. Moreover, the slightly better cetane content and heating value of isopropanol provide an alternative option to other alcohols (ethanol and methanol) [2, 23-25]. There are various papers in the literature on the utilise of isopropanol in CI engines. Alptekin [25] investigated the comparative performance and pollutant values of diesel-isopropanol and diesel-ethanol mixtures in a CRDI engine at variable speed and loads. Isopropanol was added to the diesel at a proportion of 15% by volume. When the performance test results were examined, it was stated that the fuel consumption worsened with the addition of alcohol to the diesel. Moreover, it has been reported that higher HC, CO, and NO_x values are obtained compared to diesel fuel. Liu et al. [26] examined the impacts of isopropanol utilization as an ingredient on the performance and pollutants of a CI engine for varying spray timings. isopropanol and diesel fuel were blended and the mixtures contained 5%, 10%, 15%, and 20% alcohol. The results indicated that fuel consumption deteriorated with the increment in the isopropanol substitution ratio at the same working conditions. In addition, the in-cylinder pressure and heat release rate increased, and the combustion period was shortened. Compared to diesel fuel, NO_x and HC values increased with isopropanol substitution. On the other hand, a decline in smoke values was observed as the spraying time was significantly delayed. Hazar and Uyar [27] focused on the impacts of isopropanol-diesel mixtures on performance and pollutants at varying loads and constant speed. They prepared the isopropanol-diesel mixtures involving 2%, 8%, 12%, and 18% by volume. In the paper, it was stated that isopropanol worsened fuel consumption and NO_x pollutants. Otherwise, smoke values were reported to have improved.

USE OF ISOBUTANOL IN DIESEL ENGINES

Isobutanol has the ability to be produced from biomass or fossil sources. Since isobutanol can be produced from biomass sources, it has unlimited raw materials and can be considered as a renewable fuel. In diesel-alcohol mixtures, if there is water in the mixture, this causes phase separation problem. The water affinity of isobutanol is lower than other alcohols and it has the ability to solve this problem. In the literature, there are studies on diesel-isobutanol mixtures. Karabektas and Hosoz [28] analysed the suitability of isopropanol-diesel mixtures. In addition, the performance and pollutant properties of these mixtures were investigated. Isobutanol was used in different ratios (between 5% and 20% in 5% increments) by volume. Experiments were carried out under full load, at different engine speeds in a direct injection, naturally aspirated CI engine. When the results were examined, it was determined that the engine power decreased, and the fuel consumption

worsened with the addition of isobutanol. Similarly, the thermal efficiency worsened compared to pure diesel. Emission results showed that CO and NO_x improved, while HC emissions worsened. Gu et al. [29] focused on the impact of isobutanol and nbutanol additives on pollutants and performance parameters in a CI engine. In this paper, 100% diesel, %85 diesel+%15isobutanol, %70 diesel+%30 isobutanol, %85 diesel+%15 nbutanol, and %70 diesel+%30 nbutanol mixtures are examined. The experiment tests were practiced under constant low engine speed conditions, low and medium loads, various injection times, and different EGR rates. Experiment results showed a significant improvement in soot emissions when butanol (iso or n) was substituted for diesel.

USE OF DIETHYL ETHER IN DIESEL ENGINES

Diethyl-ether has higher vaporability than other alcohol additives and consists of an ordinary molecular structure. In addition, it is capable of rapid flammability even at low oxygen concentrations. It is colorless and odorless. The oxygen density in its content is effective in reducing soot emissions. Diethyl-ether can also be produced from biomass-based materials and has unlimited raw materials for its production. Considering all these properties, diethyl-ether is an attractive additive [30, 31]. In the literature, there are studies on the adjunct of diethyl ether to diesel. Ibrahim [32] analyzed the impact of diethyl ether admixture to enhance CI engine performance and pollutants. Diethyl ether has been added to diesel fuel at varying rates up to 15% by volume. Experiments were practiced in a single cylinder CI engine at various loads and constant speed (1500 rpm). The results indicated that the adjunct of diethyl ether generally showed an increased performance trend for all loads. Moreover, the adjunct of 15% diethyl ether has been informed to enhance the thermal efficiency by about 7%. It was stated that a similar improvement occurred in fuel consumption. Rakopoulos et al. [33] experimentally analyzed the impacts of adding 8%, 16%, and 24% diethyl ether to diesel by volume in a high-speed CI engine. The study was carried out under various loads (three different loads) and at 2000 rpm. Results: NO_x emissions have improved compared to the main fuel. Similarly, CO emissions decreased due to the oxygen concentration in diethyl ether. However, HC emissions have increased. Smoke emissions tended to decrease significantly. Apart from these, as the diethyl ether ratio in diesel fuel increased, fuel consumption increased. As a result, they reported that the utilize of diethyl ether in diesel fuel is promising in terms of emissions. Banapurmath et al. [34] found that the adjunct of diethyl ether to the main (diesel) fuel enhanced engine performance and increased NO_x emissions and reduced HC, CO, and smoke pollutants. Paul et al. [35] focused on the effect of diethyl ether additive in a diesel engine. Diethyl ether was used in two different volumetric ratios. Thermal efficiency and NO_x emissions improved with 5% diethyl ether additive by volume. However, increasing the diethyl ether ratio to 10% worsened the performance.

USE OF IBE IN DIESEL ENGINE

Nowadays, many researchers are studying the impact of isopropanol-butanol-ethanol alcohols added to primary fuel (diesel) on performance and pollutants. Fei et al. [36] worked the impacts of the adjunct of IBE to diesel and the piston coating on combustion performance and emissions. Experiments were carried out on 20% by volume mixture of IBE, pure diesel and coated-uncoated pistons at varying load conditions. The results showed that coating on the piston improved fuel consumption by 5% and 7.3% for pure diesel and IBE20 fuels, respectively. It has also been reported that the smoke opacity is significantly reduced. Lee and colleagues [37] observed that mixtures of IBE and diesel improved smoke emissions in a CI engine. Otherwise, NO_x values were reported to deteriorate. Li et al. [38] examined the influence of injection applications in an IBE-diesel engine. They used a common rail diesel engine in the tests. The IBE was set at 15% and 30% by volume, respectively, and administered under various injection strategies. The results showed that mixtures containing 30% by volume IBE performed better with the double injection application. It was also reported that lower smoke pollutants were obtained at both 15% IBE and 30% IBE mixing ratios.

In Table 1, the characteristics of the main fuel (diesel), and additives (ethanol, isopropanol, isobutanol and diethyl ether) are given. [2, 17, 28, 32, 39].

Table 1. Characteristics of the main fuel and additives

Properties/Fuels	Diesel (C_xH_y)	Ethanol (C_2H_5OH)	Isopropanol (C_3H_7OH)	Isobuthanol (C_4H_9OH)	Diethyl ether $(C_4H_{10}O)$
Mole weight	~194	46	~60	~74	~74
Content of C (weight %)	~86.1	~52	60	~65	~64.8
Content of H (weight %)	~13.9	13	~13.4	~13.5	13.5
O_2 content (weight %)	0	~35	~26.6	~21.5	~21.7
Intensity (kg/m³)	830-840	~791	~805	~810	~715
Cetane value	45-52	~9	10-12	<15	100-125
LHV (KJ/kg)	~42500	~26400	~30500	~33000	~36850

MATERIAL AND METHOD

Material

A small CI (compression ignition) engine with a stroke volume of 315 cm³, two valves (one intake, one exhaust), and direct injection, was chosen for testing different alcohol-blended fuels. In addition, the AD320 model engine is air-cooled and works according to the four-stroke principle. Other characteristics of the engine are given in table 2. The engine prepared for different alcohol mixture experiments is presented in fig. 1.

Table 2. Technical characteristics of the AD320

Characteristic	Value
Bore dia. / cm	7.8
Stroke dia. / cm	6.6
Compression ratio	17.5:1
Max engine torque / Nm	10.5
Max engine torque speed / rpm	1850
Piston bowl chamber	Re-entrant bowl

Different alcohol mixture experiments were implemented in the automotive workshop of Karamanoğlu Mehmetbey University Technical Sciences Vocational School with 50 kW engine power and 157 Nm braking torque to evaluate performance. The dynamometer to control the engine for different alcohol mixture experiments is presented in figure 2.

Figure 1. Engine used for different alcohol mixture experiments.

Engine control was provided remotely (via a computer) through local software. The measurement was carried out by adjusting the motor, the desired torque, and the speed via the software. Engine accelerator operation was carried out by means of a remote-controlled accelerator pedal.

Figure 2. Dynamometer view.

Engine fuel consumption values were evaluated with a precision balance which has a sensitivity of 1%. One-minute periods were kept for measurement. The precision balance electronic screen was photographed and recorded for each minute. Figure 3 shows the fuel consumption measurement system.
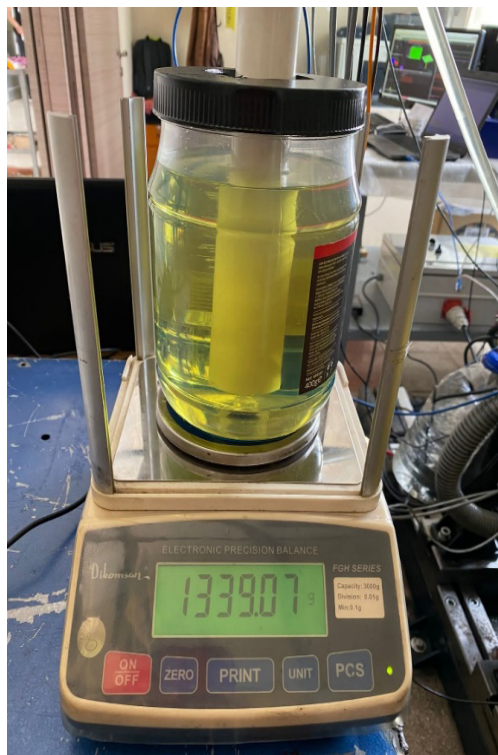
Figure 3. Fuel consumption measurement view.

NO emission and smoke opacity were measured with Bosch brand BEA70 and BEA60 model devices, respectively.

METHOD

The study was practiced in four cases. The first case was performed with diesel fuel. Then, in the second case, isopropanol, isobutanol, and ethanol alcohols were added to the diesel fuel, and their tests were conducted. In the third and fourth cases, diethyl ether was added to the diesel-alcohol fuel mixture. The fuel mixture ratios used in the experiments are given in Table 3. Performance and emission data obtained in all tests were compared with each other. Table 4 shows the experimental test matrix.

Table 3. The fuel mixture ratios used in the experiments.

Test Fuels	Volumetric ratio (%, v/v)				
	Diesel	Isopropanol	Isobutanol	Ethanol	Diethyl ether
Diesel	100	-	-	-	-
Mix 1	85	5	5	5	-
Mix 2	82.5	5	5	5	2.5
Mix 3	80	5	5	5	5

Table 4. Test matrix.

Experiment	Fuel	Load (Nm)	Engine speed (rpm)	Injection timing (CAD)	Injection pressure (MPa)
Case 1	Diesel	3 and 6 Nm	1500		
			2000		
			2500		
			3000		
Case 2	Mix-1	3 and 6 Nm	1500		
			2000		
			2500		
			3000		
Case 3	Mix-2	3 and 6 Nm	1500		
			2000		
			2500		
			3000		
Case 4	Mix-3	3 and 6 Nm	1500	340	21-22
			2000		
			2500		
			3000		

RESULTS

Diesel, IBE-Diesel, and IBE-Diesel-DEE mixtures (diesel, mix-1, mix-2, and mix-3) were performed at varying loads (3 and 6 Nm) and at varying engine speeds (1500, 2000, 2500 and 3000 rpm). Experiments continued with diesel first, then continued mix 1, mix 2, and mix 3 blends. For each experiment, the engine oil temperature was expected to be around 75 °C due to stable operating conditions.

PERFORMANCE ANALYSIS

Fig. 4 presents the variance in specific fuel consumption for certain engine loads for diesel, mix-1, mix-2, and mix-3 fuels. Generally, the lowest fuel consumption values for all fuels were obtained with diesel. It is clear that when alcohol is added to diesel, the specific fuel consumption increases. Especially at low loads, the in-cylinder temperature is low and this causes unstable combustion. Low lower calorific value and high oxygen content of alcoholic fuels can cause the in-cylinder temperature to drop further. This may result in more unstable combustion. Also, the lower calorific values of mix-1, mix-2, and mix-3 result in more fuel injection. Thus, the fuel consumption rate is increased. Fuel consumption decreased for all fuels as the engine load increased from 3 Nm to 6 Nm. High engine loads result in higher in-cylinder temperatures. Thus, the combustion develops a little more. Compared to diesel at 3 Nm load, the brake specific fuel consumption of mix-1, mix-2, and mix-3 showed an average increase of approximately 14.4%, 10.3%, and 6.8%, respectively. Compared to diesel at 6 Nm load, the brake specific fuel consumption of mix-1, mix-2, and mix-3 showed an average increase of approximately 2.3%, 2.6%, and 1.6%, respectively. The adjunct of diethyl ether appears to slightly reduce the high fuel consumption of the alcoholic mixture. This is probably owing to the high cetane number in diethyl ether.
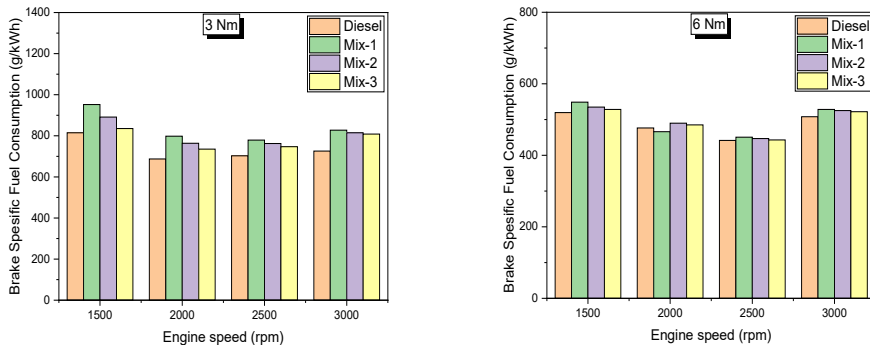
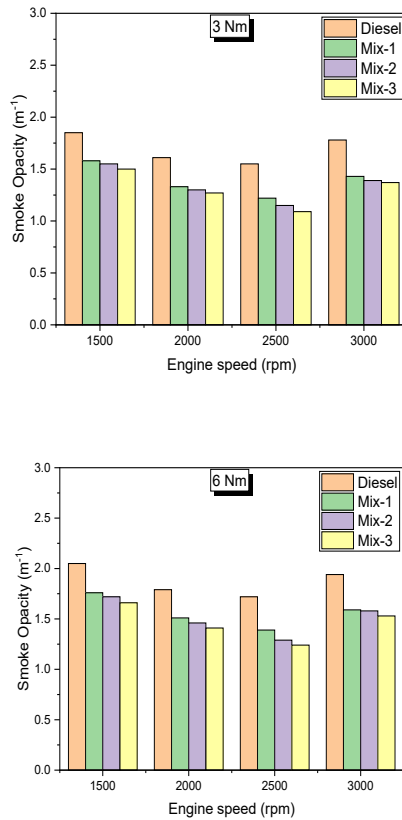**Figure 4.** BSFC variations of diesel, mix-1, mix-2, and mix-3 under varying engine speed conditions.

Figure 5 presents the variation in thermal efficiency for certain engine loads for diesel, mix-1, mix-2, and mix-3 fuels. At 3 Nm engine load, the highest thermal efficiency for diesel, mix-1, mix-2 and mix-3 fuels was obtained with diesel fuel. Improved combustion of diesel fuel compared to alcoholic blends at low loads resulted in higher thermal efficiency. Low in-cylinder temperature and high oxygen content may have caused combustion to be more unstable. It was observed that the thermal efficiency of diesel-alcohol mixtures increased slightly compared to diesel at 6 Nm engine load. The lower calorific value of alcoholic fuels may have been effective in slightly increasing the calorific value. Considering the ratio of

engine power to total fuel energy, the lower calorific value of alcohol mixtured fuels resulted in higher thermal efficiency. The highest thermal efficiency was obtained with mix-3 fuel at 6 Nm engine load. High cetane number can be considered to be effective on thermal efficiency as it improves ignition. Compared to diesel at 6 Nm load, the thermal efficiency of mix-1, mix-2, and mix-3 showed an average increase of approximately 0.9%, 2.35%, and 3.65%, respectively.
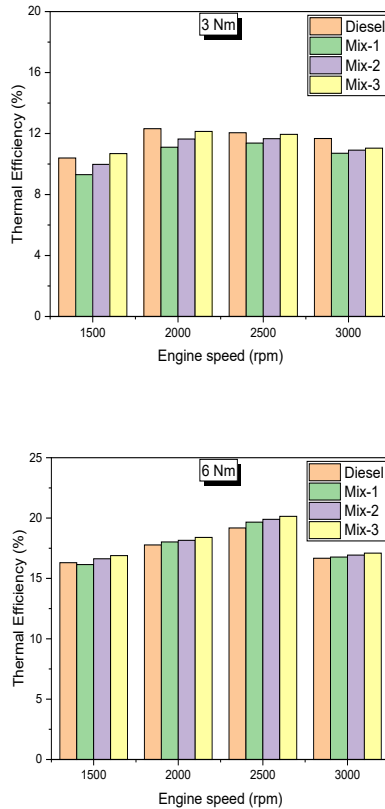
Figure 5. Thermal efficiency variations of diesel, mix-1, mix-2, and mix-3 under different engine speed conditions.

EMISSION ANALYSIS

Fig. 6 presents the variation in NO pollutants for certain engine loads for diesel, mix-1, mix-2, and mix-3 fuels. Higher engine load conditions require more fuel injection into the cylinder. This causes an increase in the in-cylinder temperature as it will provide more fuel energy. NO emissions are also likely to increase in conditions of high temperature and oxygen content [40]. As seen in Figure 6, NO emissions decreased with increasing speed. The main reason for this can be shown that complete combustion does not occur at high revs. As engine speed increases, having a shorter time to burn may reduce the likelihood of complete combustion.

The highest NO values at all speeds and load conditions were obtained with diesel fuel. The more complete combustion of diesel at high revs compared to other fuels may have increased the in-cylinder temperature. This situation resulted in higher NO values.
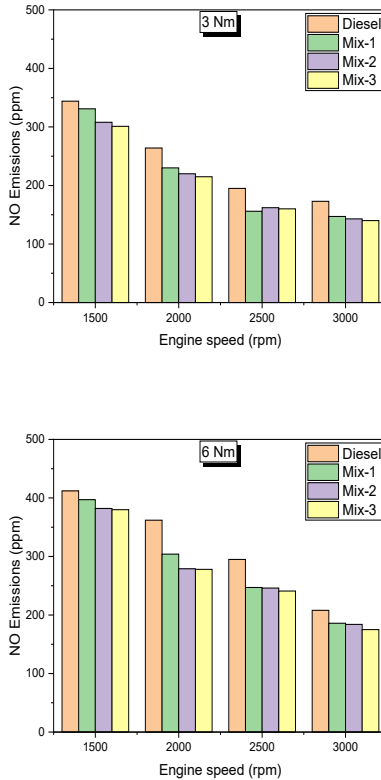
Figure 6. NO variations of diesel, mix-1, mix-2, and mix-3 under varying engine speed conditions.

Compared to diesel at 3 Nm load, the NO values of mix-1, mix-2 and mix-3 showed an average decrease of approximately 12.9%, 15.3%, and 17%, respectively. Compared to diesel at 6 Nm load, the NO values of mix-1, mix-2 and mix-3 showed an average decrease of approximately 11.6%, 14.6% and 16.3%, respectively.

Figure 7. Smoke opacity variations of diesel, mix-1, mix-2, and mix-3 under varying engine speed conditions.

Fig. 7 presents the variation in smoke pollutants for certain engine loads for diesel, mix-1, mix-2, and mix-3 fuels. The highest smoke values at all speeds and load conditions were obtained with diesel fuel. Lack of oxygen in diesel fuel and insufficient oxygen level for combustion in rich mixture zones may have increased smoke emissions. The oxygen content of alcoholic fuels can be effective in improving smoke emissions. Thanks to the higher oxygen level of Mix-3, less smoke emission has been achieved. Compared to diesel at 3 Nm load, the smoke values of mix-1, mix-2 and mix-3 showed an average decrease of approximately 18.24%, 20.8%, and 23.2%, respectively. Compared to diesel at 6 Nm load, the smoke values of mix-1, mix-2 and mix-3 showed an average decrease of approximately 16.7%, 19.5%, and 22.3%, respectively.

REFERENCES

[1] H. E. Gülcan, N. Gültekin, and M. Ciniviz, "The Effect of Methanol-Dodecanol Addition on Performance and Smoke Emission in a CI Engine with Diesel Fuel," *International Journal of Automotive Science And Technology,* vol. 6, no. 2, pp. 207-213.

[2] B. R. Kumar and S. Saravanan, "Use of higher alcohol biofuels in diesel engines: A review," *Renewable and Sustainable Energy Reviews,* vol. 60, pp. 84-115, 2016.

[3] L. Benbrahim-Tallaa *et al.*, "Carcinogenicity of diesel-engine and gasoline-engine exhausts and some nitroarenes," *The lancet oncology,* vol. 13, no. 7, pp. 663-664, 2012.

[4] L. V. Giles, C. Carlsten, and M. S. Koehle, "The effect of pre-exercise diesel exhaust exposure on cycling performance and cardio-respiratory variables," *Inhalation toxicology,* vol. 24, no. 12, pp. 783-789, 2012.

[5] W. B. Innes, "Effect of nitrogen oxide emissions on ozone levels in metropolitan regions," *Environmental science & technology,* vol. 15, no. 8, pp. 904-912, 1981.

[6] A. Mills and S. Elouali, "The nitric oxide ISO photocatalytic reactor system: measurement of NOx removal activity and capacity," *Journal of Photochemistry and Photobiology A: Chemistry,* vol. 305, pp. 29-36, 2015.

[7] J. Devaraj, Y. Robinson, and P. Ganapathi, "Experimental investigation of performance, emission and combustion characteristics of waste plastic pyrolysis oil blended with diethyl ether used as fuel for diesel engine," *Energy,* vol. 85, pp. 304-309, 2015.

[8] J. Lane, "Biofuels mandates around the world: 2016," *Biofuels Digest,* vol. 3, 2016.

[9] M. El-Adawy, A. Ibrahim, and M. El-Kassaby, "An experimental evaluation of using waste cooking oil biodiesel in a diesel engine," *Energy Technology,* vol. 1, no. 12, pp. 726-734, 2013.

[10] A. Ibrahim, M. El-Adawy, and M. El-Kassaby, "The impact of changing the compression ratio on the performance of an engine fueled by biodiesel blends," *Energy Technology,* vol. 1, no. 7, pp. 395-404, 2013.

[11] A. C. Hansen, Q. Zhang, and P. W. Lyne, "Ethanol–diesel fuel blends—a review," *Bioresource technology,* vol. 96, no. 3, pp. 277-285, 2005.

[12] N. Kosaric and J. Velikonja, "Liquid and gaseous fuels from biotechnology: challenge and opportunities," *FEMS Microbiology Reviews,* vol. 16, no. 2-3, pp. 111-142, 1995.

[13] S. Prasad, A. Singh, and H. Joshi, "Ethanol as an alternative fuel from agricultural, industrial and urban residues," *Resources, Conservation and Recycling,* vol. 50, no. 1, pp. 1-39, 2007.

[14] K. Owen and T. Coley, "Automotive fuels reference book," 1995.

[15] E. Torres-Jimenez, M. Svoljšak-Jerman, A. Gregorc, I. Lisec, M. P. Dorado, and B. Kegl, "Physical and chemical properties of ethanol− biodiesel blends for diesel engines," *Energy & fuels,* vol. 24, no. 3, pp. 2002-2009, 2010.

[16] P. Kwanchareon, A. Luengnaruemitchai, and S. Jai-In, "Solubility of a diesel–biodiesel–ethanol blend, its fuel properties, and its emission characteristics from diesel engine," *Fuel,* vol. 86, no. 7-8, pp. 1053-1061, 2007.

[17] J. Huang, Y. Wang, S. Li, A. P. Roskilly, H. Yu, and H. Li, "Experimental investigation on the performance and emissions of a diesel engine fuelled with ethanol–diesel blends," *Applied Thermal Engineering,* vol. 29, no. 11-12, pp. 2484-2490, 2009.

[18] I. Ahmed, "Oxygenated diesel: emissions and performance characteristics of ethanol-diesel blends in CI engines," *SAE technical paper,* vol. 2001, 2001.

[19] L. Xing-Cai, Y. Jian-Guang, Z. Wu-Gao, and H. Zhen, "Effect of cetane number improver on heat release rate and emissions of high speed diesel engine fueled with ethanol–diesel blend fuel," *Fuel,* vol. 83, no. 14-15, pp. 2013-2020, 2004.

[20] A. Bilgin, O. Durgun, and Z. Sahin, "The effects of diesel-ethanol blends on diesel engine performance," *Energy sources,* vol. 24, no. 5, pp. 431-440, 2002.

[21] B.-Q. He, S.-J. Shuai, J.-X. Wang, and H. He, "The effect of ethanol blended diesel fuels on emissions from a diesel engine," *Atmospheric Environment,* vol. 37, no. 35, pp. 4965-4971, 2003.

[22] R.-D. Zhang, H. He, X.-Y. Shi, C.-B. Zhang, B.-Q. He, and J.-X. Wang, "Preparation and emission characteristics of ethanol-diesel fuel blends," *Journal of Environmental Sciences,* vol. 16, no. 5, pp. 793-796, 2004.

[23] M. Lapuerta, R. García-Contreras, J. Campos-Fernández, and M. P. Dorado, "Stability, lubricity, viscosity, and cold-flow properties of alcohol− diesel blends," *Energy & fuels,* vol. 24, no. 8, pp. 4497-4502, 2010.

[24] N. Yilmaz, E. Ileri, and A. Atmanli, "Performance of biodiesel/higher alcohols blends in a diesel engine," *International Journal of Energy Research,* vol. 40, no. 8, pp. 1134-1143, 2016.

[25] E. Alptekin, "Evaluation of ethanol and isopropanol as additives with diesel fuel in a CRDI diesel engine," *Fuel,* vol. 205, pp. 161-172, 2017.

[26] Y. Liu, B. Xu, J. Jia, J. Wu, W. Shang, and Z. Ma, "Effect of injection timing on performance and emissions of di-diesel engine fueled with isopropanol," in *2015 International Conference on Electrical, Electronics and Mechatronics*, 2015: Atlantis Press, pp. 133-137.

[27] H. Hazar and U. Mahmut, "Experimental investigation of isopropyl alcohol (IPA)/diesel blends in a diesel engine for improved exhaust emissions/Egzoz Emisyonlarının İyileştirilmesi için Izopropil Alkol (IPA)/Dizel Karışımlarının Bir Dizel Motorda Deneysel İncelenmesi," *International Journal of Automotive Engineering and Technologies,* vol. 4, no. 1, pp. 1-6, 2015.

[28] M. Karabektas and M. Hosoz, "Performance and emission characteristics of a diesel engine using isobutanol–diesel fuel blends," *Renewable Energy,* vol. 34, no. 6, pp. 1554-1559, 2009.

[29] X. Gu, G. Li, X. Jiang, Z. Huang, and C.-f. Lee, "Experimental study on the performance of and emissions from a low-speed light-duty diesel engine fueled with n-butanol–diesel and isobutanol–diesel blends," *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering,* vol. 227, no. 2, pp. 261-271, 2013.

[30] N. Vin, O. Herbinet, and F. Battin-Leclerc, "Diethyl ether pyrolysis study in a jet-stirred reactor," *Journal of Analytical and Applied Pyrolysis,* vol. 121, pp. 173-176, 2016.

[31] Q. Jing *et al.*, "Deflagration evolution characteristic and chemical reaction kinetic mechanism of JP-10/DEE mixed fuel in a large-scale tube," *Fuel,* vol. 322, p. 124238, 2022.

[32] A. Ibrahim, "Investigating the effect of using diethyl ether as a fuel additive on diesel engine performance and combustion," *Applied Thermal Engineering,* vol. 107, pp. 853-862, 2016.

[33] D. C. Rakopoulos, C. D. Rakopoulos, E. G. Giakoumis, and A. M. Dimaratos, "Characteristics of performance and emissions in high-speed direct injection diesel engine fueled with diethyl ether/ diesel fuel blends," *Energy,* vol. 43, no. 1, pp. 214-224, 2012.

[34] N. R. Banapurmath, S. V. Khandal, R. L. Swamy, and T. K. Chandrashekar, "Alcohol (ethanol and diethyl ethyl ether)-diesel blended fuels for diesel engine applications-a feasible solution," *Adv. Automob. Eng.,* p. 117, 2015.

[35] A. Paul, P. K. Bose, R. Panua, and D. Debroy, "Study of performance and emission characteristics of a single cylinder CI engine using diethyl ether and ethanol blends," *Journal of the energy institute,* vol. 88, no. 1, pp. 1-10, 2015.

[36] C. Fei *et al.*, "Combustion and emission performance of isopropanol-butanol-ethanol (IBE) mixed with diesel fuel on marine diesel engine with nano YSZ thermal barrier coating," *Energy,* vol. 256, p. 124683, 2022.

[37] T. H. Lee, H. Wu, A. Hansen, T. Lee, and G. Li, "Comparison Study on Combustion and Emission Characteristics of ABE/IBE-Diesel Blends in a Common-Rreail Diesel Engine," SAE Technical Paper, 0148-7191, 2017.

[38] G. Li, T. H. Lee, Z. Liu, C. F. Lee, and C. Zhang, "Effects of injection strategies on combustion and emission characteristics of a common-rail diesel engine fueled with isopropanol-butanol-ethanol and diesel blends," *Renewable Energy,* vol. 130, pp. 677-686, 2019.

[39] H. E. Gülcan, M. Ciniviz, and N. Gültekin, "Assessment of the influence of propanol-diesel and butanol-diesel blends on performance and emissions in a small diesel engine," in *1st International Conference on Innovative Academic Studies*, Konya/Turkey, 2022: ICIAS, pp. 614-618.

[40] W. W. Pulkrabek, "Engineering fundamentals of the internal combustion engine," ed, 2004.

INVESTIGATION OF TRIBOLOGICAL BEHAVIOR FOR SUNFLOWER OIL AS A BIOLUBRICANT IN INTERNAL COMBUSTION ENGINE-REVIEW STUDY

Omar AL-HADEETHI[1], A. Engin ÖZÇELİK[2], M. Turan DEMİRCİ[3]

INTRODUCTION

Several researchers around the world aims to optimize the environmental and enhances machines efficiency by utilizing lightweight materials, recirculation exhaust gases, uses less harmful fuel to optimize the combustion quality. Energy consumption in worldwide increasing significantly day by day, Figure 1 demonstrate Energy consumption in worldwide. Fossil fuel considered main Energy sources and burning it with large quantities can be contribute in earth warming because it emits harmful gases [1] and increased the possibility for fossil fuel depletion[2]. Several studies have reported an energy consumption rate of about 53% by 2030 generating concerns in governments [3].
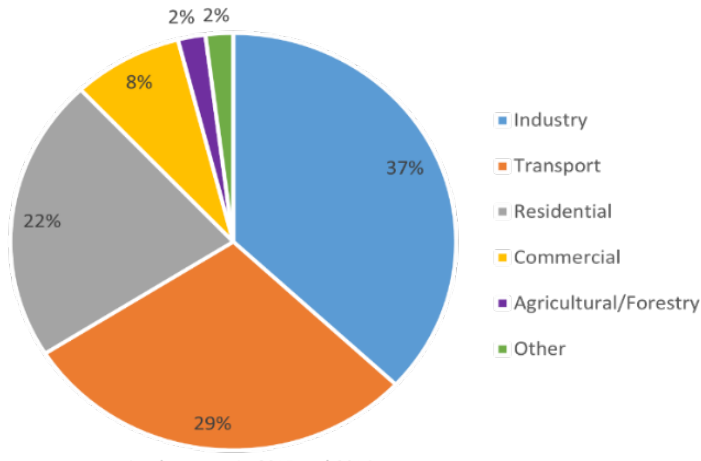
Figure 1: Energy consumption in worldwide [4]

Moreover, emissions produced from fossil fuels through burning have damaging effect on the environment, and living things [4-5].

Several researchers worldwide aim to optimize the environment and enhance machines' efficiency by utilizing lightweight materials, recirculation exhaust gases, and less harmful fuel to optimize combustion quality. Lubrication process defines as the process that employed lubricants fluid between two friction surfaces. Major

1   Selcuk University,Technology Faculty,Mechanical Engineering Department, Konya, TURKEY
2   Selcuk University,Technology Faculty, Mechanical Engineering Department, Konya, TURKEY
3   Selcuk University,Technology Faculty,Mechanical Engineering Department, Konya, TURKEY

role of the lubricant fluid is to reduce the friction force and wear between two moving parts. The main function of the lubrications process is prevented heat losses produced due to contact moving parts, reduce corrosion, reduce, acts as an insulator in the electric transformer, and, lubricant fluid contributes to dust, dirt cleaning. For a long period, traditional oils derived from fossil fuel has been utilized in automotive engine, but their limitation by low availability of fossil fuel, moreover using traditional oils derived from fossil fuel has a damaging effect on the environment and human health. fossil fuel depletion increases fossil fuel prices and improves the environmental quality, all these concerns pay to innovate worldwide to search for alternative solutions to reduce environmental degradation and increase fossil fuel price concerns. Bio lubricants, alternative, renewable, sustainable, non-edible, non-toxic. Bio lubricants may be gradually replaced traditional lubricant derived from fossil fuel in near future. Figure 3 (a) showing the Geographical representation of-Sunflower growing regions. Figure 3 (b) showed virous kind of bio lubricant consumption in the worldwide [6-7-8]. Bio lubricants derived from vegetable oil characterized by high lubricity properties, rise in the flash point, low volatility, and rise in the viscosity. Long chain of fatty acid and polar compounds in the chemical structure made bio lubricants a proper alternative solution to traditional oil due to high lubricity properties. Table 1 describing a chemical composition of fatty acid of sunflower oil. In the last decades, selection of lubrication or fuel or any fluid used in the manufactory proses is based on their costs and doesn't pay any attention to environmental issues. But nowadays, environmental issues, become the most significant concerns worldwide, on the other hand, increases fossil fuel prices and increase the possibility of fossil fuel depletion. All these reasons paid researchers to develop bio alternative sources to decrease the dependence on fossil fuel, bio lubricants can be obtained from edible and non-edible oil sources such as (jojoba, soybean, canola, palm, rapeseed, Sunflower oilseed, and jatropha, etc.). Sunflower oilseeds are renewable plants. Sunflower oilseeds consist of saturated fatty acid with portion 15%, and unsaturated fatty acid with portion 85%. Sunflower oil grows in various kinds of soil and not needed special equipment's Sunflower oil contains a red flower, or orange flower as can be snowed in Figure 2.

Sunflower oil extracted by pressing on Sunflower seeds. Sunflower oil considered an important sources feedstock for vegetable oil in the worldwide [10]. Main challenge toward sunflower oil marketing is low thermal stability; these features depend on the unsaturated fatty acid concertation. In last decades, farmers succeed in overcome on the low thermal stability problems by Sunflower oil chemical modification. Unique properties of Sunflower oil made them a desirable kind of vegetable oil to operate efficiently in internal combustion engine as a lubricant due to a long chain of fatty acids that is combined with the material surface to produce a thin layer (metallic soap layer). Sunflower oil have a good anti- wear properties than conventional oil [11]. Sunflower oil has importance in manufactory

and automotive usage. Sunflower oil has several benefits like availability, low cost, renewable and sustainability, eco-friendly, biodegradable, zero emissions, and high resistance to various diseases [12]. Sunflower oil are drought and resist for salt plant [11]. Table 2 summarizes the Physical characteristics of Sunflower oil. This study reviews the physicochemical properties of Sunflower oil, also summarize the chemicals modifications as adding Nanoparticle to Sunflowers oil, and investigated tribological characteristics of Sunflower oil when using it in internal combustions engine as Bio lubricant to reduce friction and wear between moving parts. Desirable properties of Sunflower oil are affected on wear rate in internal combustion engine by various factors such as followings

- Wear in the engine's parts (Lubrication, Viscosity, Carbon deposed.
- Exhaust gases quality, easy operation, quality of ignition, Calorific value)
- Low-temperature properties (Pour point, Cloud point)
- Properties including storage and transport (Oxidative stability, Flash point).

Figure 2: Sunflower plant [13]

Table 1. Chemical composition of fatty acid of various Bio lubricants [14]

Fatty acid %	Palm oil	Soybean oil	Sunflower oil
Myristic acid [C 14:0]	1.5	-	-
Palmitic acid [C 16:0]	43	11	2.5
Palmitoleic acid C 16:1	-	Trace	Trace
Stearic acid [C 18:0]	5	4.5	2
Oleic acid [C 18:1]	40	23	89.5
Linoleic acid [C 18:2]	10	55	6
linolenic acid [C 18:3]	-	Trace	0
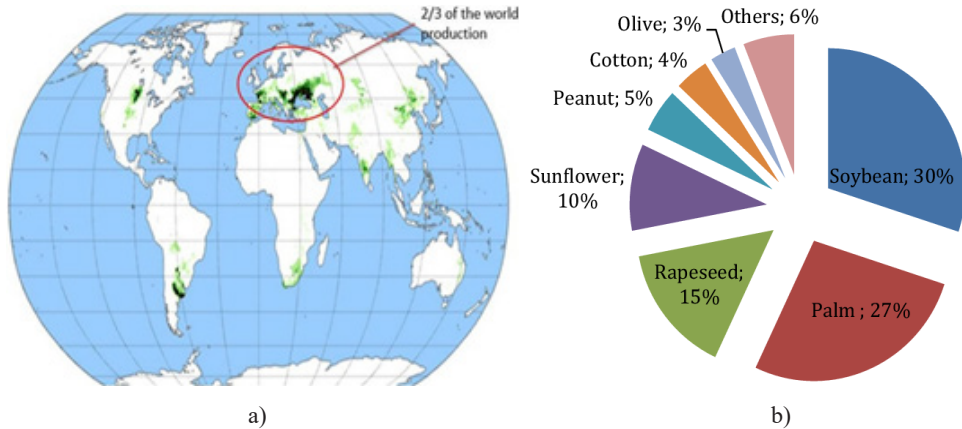
Figure 3 a): Geographical representation of-Sunflower growing regions [14].b): Various kind of bio lubricant consumption in the worldwide [15]

PHYSICAL CHARACTERISTICS OF SUNFLOWER OIL

The physical characteristics obtained from sunflower seed oils are affected by many aspects, including the status of raw materials, chemical composition of sunflower seeds, long fatty acids chain present in sunflower oil, manner in which sunflower oils are refined, and the method of production of sunflower oils.

Oxidative stability: Oxidative stability define as resistance of edible oil to oxidation during storage. Quality and effectiveness of lubricants affected strongly by Oxidative stability. There are two main kinds of lubricants oxidation, autoxidation and photosensitized oxidation responsible to lubricants oxidation. Oxidation takes places through chemical reaction between fatty acid methyl ester and air resulting in Biolubricants oxidization. Biolubricants acidification can be affected the bio-lubricants quality. Chemical changing resistance to store which take placing for bio lubricant during store through long time. Main reasons for the bio lubricant poor stability and oxifacation are the exposed to air and sunlight directly. Dual bonds of the unsaturated fatty acids for methyl ester exposed to acidification. In reality, structure of fatty acid evaluated the oxidative stability of sunflower oil. Sunflower oil rich to polyunsaturated fatty acid (PUFAs) in chemical structure has a lower oxidative stability than sunflower oil contains monounsaturated fatty acid (MUFAs). Oxidative stability of sunflower oil decreased with increasing in double bonds of the carbon in chemical structure. Adding anti-oxidifacation for bio lubricant to improve oxidative stability. Generally, oxidative stability of the Sunflower oil is 8 hours based on the EN14214 standard. In internal combustion engine, through the combustion in engine, engine lubricating oil forms a layer to prevent wear between engine parts, for this reason, engine oil undergoes physical and chemical changing resulting in oxidation in oil [13-14-15-16].

Viscosity: high viscosity of bio lubricants is considered the main problem in the use of bio lubricants as pure oils in an automobile engine, for these reasons, large

number of researchers have made efforts to reduce the viscosity of biolubricants to develop the features of biolubricants and found that esterification reactions is the best way to decrease the viscosity of biolubricants. Briefly, esterification reactions are chemical reaction between carboxylic acid (RCOOH) with alcohol (ROH) to produce ester (RCOOR) and water. As can be shown in the Figure 5. Viscosity of sunflower oil is decreased about five times through esterification reactions [20]. Esterification reaction are most commonly used way to reducing biolubricants viscosity.
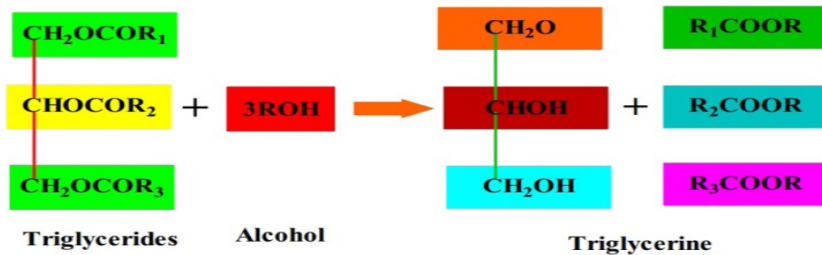
Figure 5: Esterification chemical reaction [21]

Ester content: feature of Ester content for Sunflower oil determines the purity of Sunflower oil, this feature maybe change based on the feedstock of Sunflower oil production. Gas chromatography (GC) devise mostly utilized to evaluate the ester content for Sunflower oil. In the Sunflower oil ester contain must be 96.5% based on the EN14214 standard. Sunflower oil ester contain lower than 96% maybe result in undesirable reaction conditions or form different Contamination in Sunflower oil, sterol, alcohols, glyceride, and non-decomposable glycerol, all these unwanted results can be removed by distillation process [13-14].

Viscosity index: viscosity index of bio lubricant evaluated the viscosity change with temperature differences. always measured from viscosity measured at 40 °C or 100 °C. Increasing in viscosity index indicates to reducing in temperature impact.

Flash point: Flash point is the low temperature of the lubricants measured at a pressure of 1 bar, which is equivalent to 101.3 KPa, as it causes the evaporation of the oil. The flash point determines the risks as it limits the flammability.

Pour point: The pour point is the low temperature at which the biollubricant flows hardly, especially under cold climates operating conditions. Viscosity of Biolubricants affected strongly by temperatures and pressures Figure 9 shows the variation of sunflower oil and mineral oil with Temperature [22]. (Lemuel M. Diamante et al 2014) experimental studies the effect of higher shear rate (64 to 4835 s^{-1}) on absolute viscosities of different vegetable oil at different temperature (26°C to 90°C), using a Lamy Viscometer RM100, rotating viscometer with C-coaxial cylinder. Results have been showed, all the vegetable oil have a characteristic of

Newtonian fluid. Race bran oil was highly viscos between other kinds of oil (0.0389 Pa.s at 38 ℃), Walnut oil was a lower viscos (0.0296 Pa.s at 38 ℃). Kinematic viscosity of Sunflower oil at 40 ℃ (32.50 mm^2/s) also the absolute viscosity of different vegetable oil kinds decreases with increasing temperature. Sunflower oil have a highest activation energy. Table 4 chemical modification of sunflower oil ways [23].

Density: density of any fluids can be defined as mass substance to unit volume, density is an important property which is limited the qualities of lubricant, density of lubricants aground 0.85-0.90 g/cm^3 at the 15℃ temperatures based on EN14214 standard [24]. Density of Sunflower oil affected by various aspects like quantities of fatty acid, purity of oil, and, feedstock structures. Generally, density of bio-lubricants considered higher than density of lubricant derived from fossil fuel. Density of Sunflower oil can be increasing by reduced the fatty acid chain long, and increases double bonds, on other hand, the high lubricants density means high Lubricants energy contained and can be resistance to wear efficiently [18]. High density of sunflower oil is an important indication to high glycerol contains and cannot remove it completely through esterification reaction. Adding nanoparticle to sunflower oil can be improve the tribological properties of bio lubricants [25]. (Vicente Cortes et al, 2020) experimentally studied the effect of adding Silicon dioxide SiO_2, and Titanium dioxide TiO_2 nanoparticles as additives to sunflowers oil, on tribological characteristics and rheological properties. Parallels plat "rheometer" used to determine the effect of concentrations and shear rate on the shear viscosity. experiments results showed the coefficient of friction decrease with additions silicon dioxide SiO_2 about 77%, also, coefficient of friction decrease with additions Titanium dioxide TiO_2 to sunflowers oil about 93% [38].(Abdulmunem R.Abdulmunem, et al 2018) experimentally investigated the effects of depression of MWCNT on sunflower oil and using it as lubricant with three concentration ratio 0.1, 0.2, 0.3%, using in experiments four stroke single cylinder diesel engine under different test conditions engine speed, and engine load.

Results showed the using non-edible sunflowers oil can be reduces the specific fuel consumption 6%, and reduces cylinders surfaces temperatures of internal combustion engine about 4%, and increases in brake thermal efficiencies of engines 7%. Adding of MWCNT on sunflower oil shows an improvement in engines performances and reduction in specific fuel consumption around (10,13,17%), also showed reduction in cylinders surface temperatures (8, 12, 16%) [25-26].

Table 2 : Physical-Chemical properties of sunflower oil [27]

Properties	Values
Iodine value (g of I_2/100g)	142.3
Acid value (mg KOH/g)	0.22
Saponification value (mg KOH/g)	16.69
Flash point (℃)	212
Pour point (℃)	11
Density at 15 ℃ (kg/m3)	923
Kinematic viscosity 40 ℃ (mm²/s)	32.50
Viscosity index	100
Cloud point ℃	-4
Water content (mg/kg)	417.69
Pour point ℃	-18
Boiling point (℃)	300-600
PH	5

SUNFLOWER OIL APPLICATIONS

Biolubricants are derived from edible and non-edible oil like (jojoba oil, soybean oil, canola oil, palm oil, rapeseed oil, Sunflower oil, and jatropha oil). Sunflower oil extracted by pressing on Sunflower seeds. Sunflower seeds renewable plants, Sunflower oil consist of saturated fatty acid 15%, and unsaturated fatty acid 85%. Sunflower oil grows in various kinds of soil and not needed special equipment's Sunflower oil contains a red flower, or orange flowers.

Sunflower oil considered an important sources feedstock for vegetable oil in the worldwide [10]. Main challenge toward sunflower oil marketing is low thermal stability; these features depend on the unsaturated fatty acid concertation. In last decades, farmers succeed in overcome on the low thermal stability problems by Sunflower oil chemical modification. Unique characteristics of Sunflower oil made them a desirable kind of vegetable oil to operate efficiently in internal combustion engine as a lubricant due to a long chain of fatty acids that is combined with the material surface to produce a thin layer (metallic soap layer), Sunflower oil have a good anti- wear properties than conventional oil [11]. Superior characteristics of sunflower oil as low toxicity and environmental benefits made them have a wide applications of bio lubricants in the automotive. bio lubricants applications in the automotive in various ways such as gear oil, wheel bearing grease, brake fluid, string fluid, engine oil, transmission fluid, gearbox oil, lubricant oil for two-stroke engines, hydraulic turbine fluid, grease, and compressor oil as can be seen in Figure 6.

Bio Lock Oil
Pour Point: -30°C
Flash Point: 295°F
Viscosity: 180VI

Bio Gear Oil
Pour Point: -30°C
Flash Point: 536°F
Viscosity: 180VI

Bio Transmission Oil
Pour Point: -20°C
Flash Point: 600°F
Viscosity: 183VI

Bio Power Biodiesel
Pour Point: 32°F
Flash Point: 220°F
Lowest Emission
Improved Fuel Economy

Bio Bearing Grease
Pour Point: -36°C
Flash Point: 536°F
Viscosity: 186VI

Bio Hydraulic Fluids
Pour Point: -36°C
Flash Point: 532°F
Viscosity: 194VI

Bio Engine/Motor Oil
Pour Point: -18°C
Flash Point: 419°F
Viscosity: 220VI

Bio Steering Fluid
Pour Point: -36°C
Flash Point: 520°F
Viscosity: 190VI

Bio Wheel Bearing Grease
Pour Point: -36°C
Flash Point: 536°F
Viscosity: 186VI

Bio Brake Fluid
Pour Point: -30°C
Flash Point: 525°F
Viscosity: 190VI

Bio Chassis Grease
Pour Point: -36°C
Flash Point: 536°F
Viscosity: 186VI

Figure 6 : Several applications of bio lubricants in the automotive [30]

SUNFLOWER OIL ADVANTAGES AND DISADVANTAGES

Sunflower oil contains 10% of the oxygen that makes Sunflower oil with proper tribological properties. Sunflower oil uses as bio lubricant can reduce harmful emissions to 78% compared with traditional oil. Sunflower oil is renewable, biodegradable, non-toxic, non-flammable, eco-friendly, with a high flash point, and low volatility, high viscosity index and doesn't contain sulfur and aromatic contain these features make it an ideal lubricant. Sunflower oil reduces particle matter contains in the environment, therefore it contributed to air pollution. Sunflower oil production can be performed easily, cheapest, and doesn't require special equipment. Sunflower oil has superior lubricity properties in internal combustion engines that reduced engine wear thereby improving the internal combustion engine's efficiency, Sunflower oil has the ability to combine with nanoparticle additives. Main drawbacks of Sunflower oil used as lubricating internal combustion Engines are low thermo-oxidative stability and high freezing point in spite of sunflower oil derived from vegetable oil, have some demerits like high raw material cost, low cold flow properties through working under cold climate conditions, poor evaporation, low pour point, very long chain of fatty acid leading to lower temperature characteristics , excessive long fatty acid leading to lower in oxidative stability, various studies have been reported that transesterification reaction most proper solution to optimizing the low oxidate stability of sunflower oil. Also, to produce more quantities of sunflower oil there is a need to enhance the sunflower oil viscosity. oil viscosity is considered the main factor in determining the coefficient of friction between two moving parts as its works to form a thin protective film between two friction surfaces. Table 5 sunflower oil advantages and their benefits [29-30].

Table 5: Sunflower oil advantages and their benefits [29-30]

Sunflower oil advantages	Application
High lubricity	Reduces friction losses thereby saving energy from 5% to 15%
High viscosity index	More stability at high operation temperature 250 °C and above
Sunflower oil have low volatility	Reduces exhaust gases emission
Sunflower oil have high flash point	These features provide more safety during storage
Sunflower oil have oil mist reduction	These features provide less inhalation of vapor
Quick biodegradable	Reduces the toxicity in the exhaust emissions gases
Sunflower oil have high detergency	Remove the use of detergent as lubricant additive
High boiling temperature	Reduces exhaust emissions gases
Low Sulphur content	Decrease the toxity in the exhaust gases
Sunflower oil have high heat content	Improve combustion efficiency

TRIBOLOGICAL BEHAVIOR OF SUNFLOWER OILS IN INTERNAL COMBUSTION ENGINE

Sunflower oils have higher tribological characteristics than conventional oils, reason beyond that, sunflower oils contain long chains of unsaturated fatty acids, they combine with metal surfaces to form a thin layer (metallic soap layer). Various researchers carried out an experiment on sunflower oils as lubrication in internal combustion engines. (Abdulmunem et al, 2019) studied a difference between the conventional oil and sunflower in internal combustion engines as lubrication, for various properties, like the difference between the conventional oil and sunflower in coefficient of (coefficient of friction and engine load, engine torque and brake thermal efficiency, and lubricant viscosity and operation temperature, Variation of cylinder temperature with time). They conducted: When the loads are increased, the coefficient of friction increases as shown in Figure 13 that means more loads generated more mechanical power losses, as shown in the equations

(1-2-3-4-5) the change in the coefficient of friction with the applied loads can be calculated easily [28]. In the current century, dispersing a nanoparticle in vegetable oils is considered a new pathway to improve the tribological behavior to rise the efficiency of the machine as an internal combustion engine. several studies have reported that dispersing a nanoparticle in vegetable oils improves the tribological behavior of the internal combustion engine as can be seen in Table 6, tribological performance of Nanoparticle dispersed in bio lubricants have four main mechanisms as can be seen in Figure 7 [30-31-32-33]. Rolling effect where small spherical nanoparticle roll between friction surface material to change sliding friction into rolling friction. Polishing effect where nanoparticle contribute in surface polishing thereby reduce the friction surface material roughness.
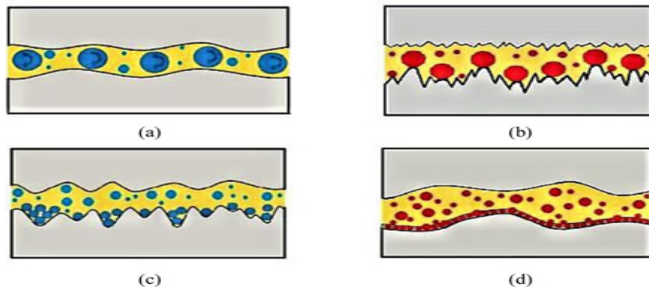
Figure 7 : Four main mechanisms to improve the tribological performance of bio lubricant through dispersing nanoparticle in bio lubricant a) rolling effect, b) polish effect, c) mending effect, d) protective film formation [30-31-32-33].

(Le Gong et al, 2015) experimentally studied the tribological characteristics of Graphite Nanoparticles dispersed in sunflower oil with varying concentrations. Experiment study investigated the tribological characteristics using a pin-on-disk Tribometer to determine coefficient of friction and wear rate. Results showed the tribological characteristics of Graphite Nanoparticles dispersed in sunflower oil with various concentrations decreasing the coefficient of friction by 53% and decreasing the wear rate by 11%. Also noted an increase in graphite nanoparticles contributed in decrease the coefficient of friction and wear rate. Rational reasons for these improvements in the tribological characteristics of Graphite Nanoparticles dispersed in sunflower oil are Graphite Nanoparticles able to form the physical oil film between friction surfaces as can be seen in Figure 8 [37]. (Sayeed et all, 2022) experimentally studied the tribological behavior of environmentally eco-friendly Halloysite Nanotube particles dispersed in sunflower oil with various concentrations of 1.5 wt.%. 0.05wt.%, under various pressure operation conditions (high-pressure operation conditions-low-pressure operation conditions). Experiment study investigated the tribological characteristics using a block on ring tribometer to determine the coefficient of friction and wear rate. results showed the tribological behavior of environmentally-eco-friendly halloysite nanotube particles dispersed in sunflower oil with different concentrations of 1.5 wt.% at low-pressure operation conditions decreasing coefficient of friction by 29%, wear rate by 70%. tribological behavior of environmentally-eco-friendly halloysite nanotube particles dispersed in sunflower oil with different concentrations of 0.05 wt.% at high-pressure operation conditions decreasing coefficient of friction by 55%, wear rate by 56% as can be seen in Figure 8 [38].
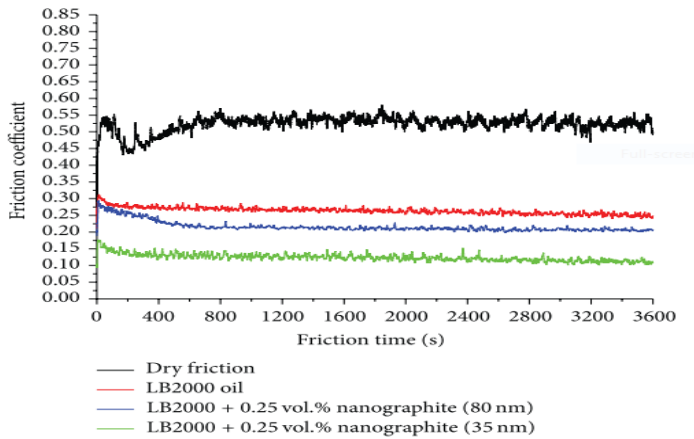
Figure 8 : Variation of friction coefficient with friction time using different lubrication [37]

(Omrani et al, 2021) experimentally studied the tribological properties of Graphite nanoparticles (GNP), dispersed in Vegetable oil with different concentrations. experiment study investigated the tribological behavior utilizing pin-on-disk tribometer to determine the coefficient of friction and wear rate. Results showed the tribological behavior of graphite nanoparticles (GNP), dispersed in Vegetable oil decreasing the coefficient of friction by 84% as can be seen in Figure 11 [39]. (Baskar et al, 2015) experimentally studied the tribological characteristics of CuO, WS_2, and TiO_2 nanoparticles dispersed in rapeseed oil with various concentrations. experiment study investigated the tribological behavior using a four-ball tribometer to determine the coefficient of friction and wear rate. results showed the tribological behavior of CuO, WS_2, and TiO_2 nanoparticles dispersed in rapeseed oil with various concentrations decreasing coefficient of friction, provide smoother wear scar, optimized tribological behavior, and higher viscosity than mineral oil [40]. (Desaria et al, 2016) experimentally studied the friction behavior of CuO, and SiO_2 nanoparticles dispersed in Soybean oil and rapeseed with various concentrations. Experiment study investigated the lubricating efficiency using a standard compression ring tester. Results showed the concentration of nanoparticles contributes to a decrease in the coefficient of friction and wear rate. results showed the tribological behavior of CuO, and SiO_2 nanoparticles dispersed in Soybean oil and rapeseed with various concentrations decreasing the coefficient of friction by 31%. Sunflower oils have a good viscosity compared to conventional oils. Using Sunflower oils as a lubricant in internal combustion engines can contribute to improving tribological characteristics and reducing friction between moving parts, Sunflower oils contribute to improving the thermal efficiency of internal combustion engines, The logical reasons beyond that when the friction between moving parts decreases, that is means the power losses will decrease, thereby results in an increase in the Brake thermal efficiency and engine torque of the internal combustion engines,

as shown by equation No. (1). Figure 13, Figure 14, Figure 15, Figure 16 shows the effect of internal combustion engine cylinder temperature during 360 minutes of continuous operation. Cylinder temperature was monitored for the first 90 minutes, where mineral oils were compared with sunflower oils. It was noticed that the cylinder temperature decreased when using sunflower oils as oils. Lubrication in internal combustion engines due to the friction between the moving parts in internal combustion engines. The effect of the operating temperature of internal combustion engines on the emission rates of carbon dioxide, carbon monoxide, nitrogen oxides, and hydrocarbons when the engine is Lubricated with Sunflower oils. Also, was noticed that increasing the speed of internal combustion engines leads to an increase in the percentage of Carbon dioxide emissions, Carbon monoxide.

$$\text{B.P} = \frac{2 \times \pi \times N \times T}{60 \times 1000} \tag{1}$$

$$\xi_{bth} = \frac{B.P}{Cv \, mf} \tag{2}$$

$$m_f = \frac{Vf \times \rho f}{tb} \tag{3}$$

$$\mu = \frac{TORQUE \times \sqrt{6}}{3 \, Wr} \tag{4}$$

$$\mathcal{M} = \text{F/N} \tag{5}$$

Where:

B.P: brake power of internal combustion engine (kW),

N: speed of internal combustion engine,

T: torque of internal combustion engine,

ξbth: brake thermal efficiency of the internal combustion engine,

 Cv: Calorific value of the fuel used in an internal combustion engine (KJ/Kg),

m_f: rate of fuel consumption in an internal combustion engine (Kg/sec),

V_f: volume of fuel (m3)

tb: time required to consumed fuel(sec),

W: load (kg), the distance between the axis of rotation and center of the contact the surface on the Lowe balls [9-24].
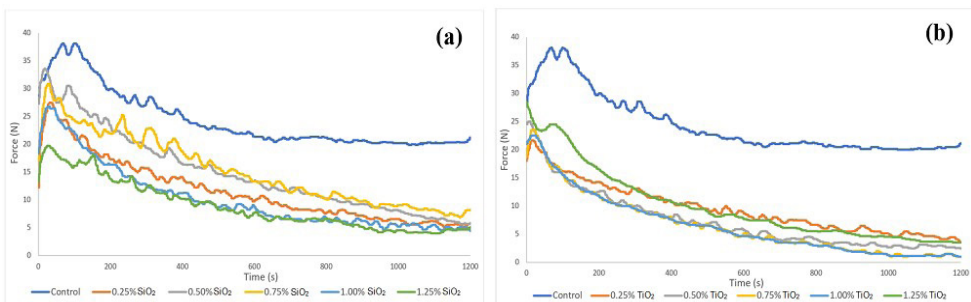
Figure 10 : Variation of friction coefficient with friction time of sunflower oil with a) (SiO2), b) (TiO2) [41]
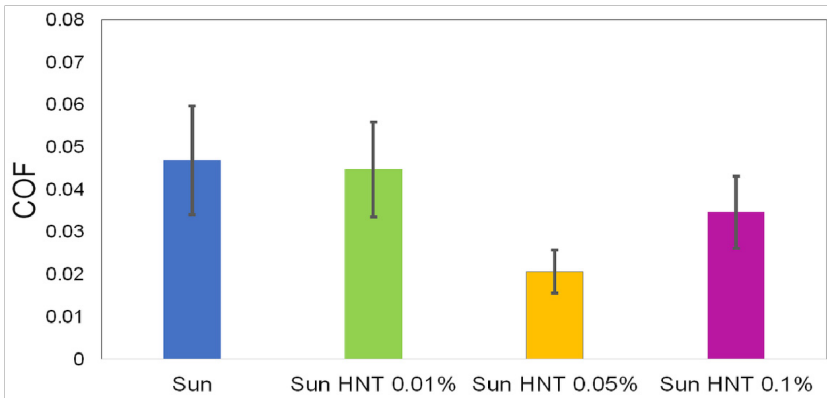
Figure 11: Variation friction coefficient with different concentration of sunflower oil and sunflower oil nanoparticle [39]
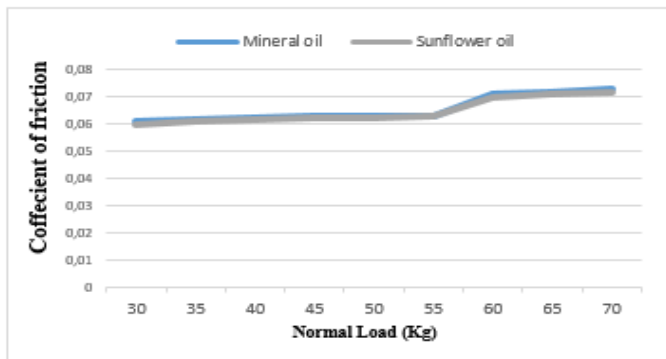
Figure 12 : Variation of friction coefficient with engine load [11]
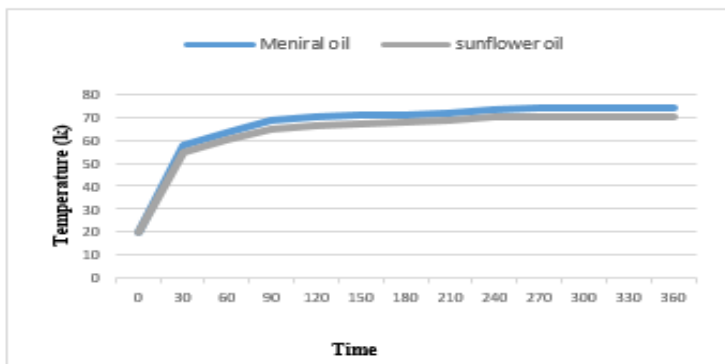
Figure 13 : Variation of cylinder temperature with time [28]
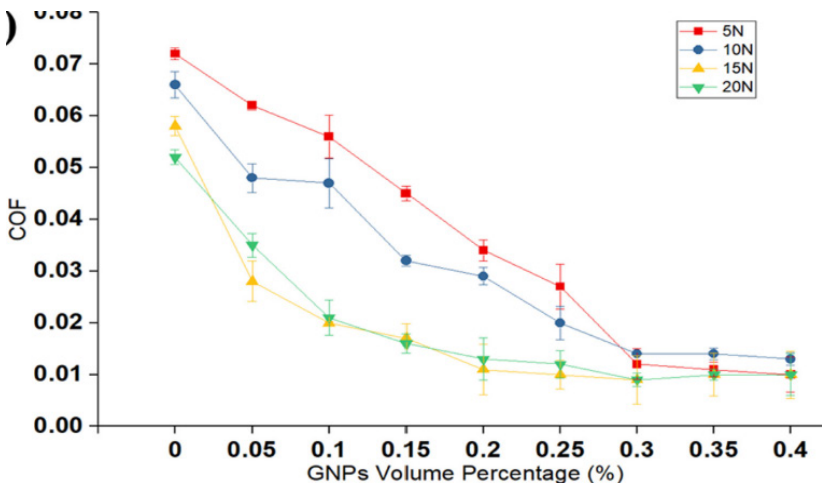
Figure 14 : Variation of friction coefficient with different graphite concertation of nanoparticle GNP [39].
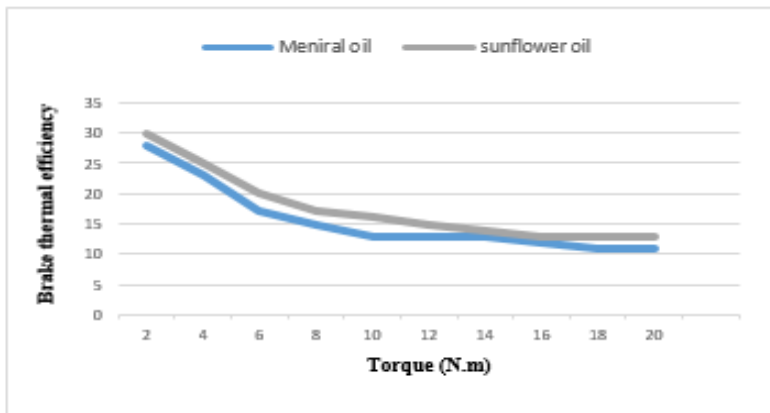
Figure 15 : Variation of Brake thermal efficiency and engine torque [11]
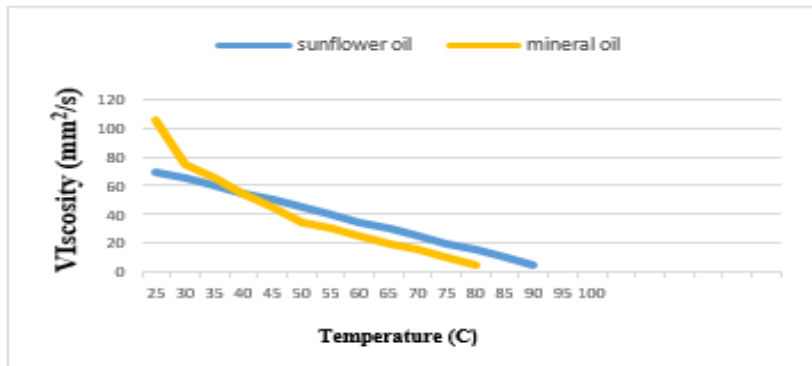
Figure 16 : Variation of lubricants viscosity and engine temperature [28]

CONCLUSION

Bio lubricants have several benefits like non-toxic, eco-friendly lubricants, zero-emissions, and degradable, Bio-lubricants considered superior's solution to replace conventional lubricants in future. This paper reviews the physicochemical characteristics of Sunflower oil, also summarize the chemicals modifications as adding Nanoparticle to Sunflowers oil, and investigated tribological behavior of Sunflower oil when using it in Internal combustions engine as Bio lubricant. Lubrication process has a vital importance in internal combustions engine because it reducing heat in moving parts of engine also it contributed in unwanted heat rejecting out of engines. the previous investigation had been founded out utilized the Sunflowers oil in internal combustion engine reduced the friction between moving parts due to long chain of fatty acids that companies with surfaces and forms a thin layer, also adding Nanoparticles as additive can be improves engines efficiency. Sunflowers oils have a several advantages to reduce the global warming thereby it contributed in environment enhancements.

REFERENCES

[1] V. Cortes, K. Sanchez, R. Gonzalez, M. Alcoutlabi, and J. A. Ortega, "as Lubricant Additives in Sunflower Oil," 2020.

[2] T. J. Wallington, C. K. Lambert, and W. C. Ruona, "Diesel vehicles and sustainable mobility in the U.S.," *Energy Policy*, 2013, vol. 54, pp. 47–53.

[3] H. C. Ong, T. M. I. Mahlia, and H. H. Masjuki, "A review on energy scenario and sustainable energy in Malaysia," *Renew. Sustain. Energy Rev.*2011, vol. 15, no. 1, pp. 639–647.

[4] R. Vooradi, S. B. Anne, A. K. Tula, M. R. Eden, and R. Gani, "Ene(1) Vooradi, R.; Anne, S. B.; Tula, A. K.; Eden, M. R.; Gani, R. Energy and CO2 Management for Chemical and Related Industries: Issues, Opportunities and Challenges. BMC Chem. Eng. 2019, 1, 7.rgy and CO2 management for chemical and related industries: ," *BMC Chem. Eng.*2019, vol. 1, no. 1, p. 7, 2019

[5] L. Lin, Z. Cunshan, S. Vittayapadung, S. Xiangqian, and D. Mingdong, "Opportunities and challenges for biodiesel fuel," *Appl. Energy*, 2011, vol. 88, no. 4, pp. 1020–1031.

[6] M. I. Arbab, H. H. Masjuki, M. Varman, M. A. Kalam, S. Imtenan, and H. Sajjad, "Fuel properties, engine performance and emission characteristic of common biodiesels as a renewable and sustainable source of fuel," *Renew. Sustain. Energy Rev.*2013, vol. 22, pp. 133–147, 2013.

[7] X. Zhang, C. Peterson, D. Reece, and R. Haws, "B b a e," 1998, vol. 5, no. Epa 410, pp. 1423–1430.

[8] A. H. Demirbas and I. Demirbas, "Importance of rural bioenergy for developing countries," *Energy Convers. Manag.*2007, vol. 48, no. 8, pp. 2386–2398.

[9] A. Yadav, Y. Singh, and P. Negi, "A review on the characterization of bio based lubricants from vegetable oils and role of nanoparticles as additives," *Mater. Today Proc.*2021, vol. 46, no. xxxx, pp. 10513–10517.

[10] M. Akgün and E. Söylemez, "Determining the Future Trends of Safflower Plant in Türkiye," *Int. J. Agric. Environ. Food Sci.*2022, vol. 6, no. 1, pp. 50–57.

[11] M. H. Jabal, A. R. Abdulmunem, and H. S. Abd, "Experimental investigation of tribological characteristics and emissions with nonedible sunflower oil as a biolubricant," *J. Air Waste Manag. Assoc.*2019, vol. 69, no. 1, pp. 109–118.

[12] H. K. Trivedi and D. V. Bhatt, "Effect of lubricating oil on tribological behaviour in pin on disc test rig," *Tribol. Ind.*2017, vol. 39, no. 1, pp. 90–99.

[13] M. H. Muller *et al.*, "Occurrence, distribution and distinctive morphological traits of weedy Helianthus annuus L. populations in Spain and France," *Genet. Resour. Crop Evol.*2009, vol. 56, no. 6, pp. 869–877.

[14] T. Inovia, "Sun fl ower in the global vegetable oil system: situation, speci fi cities and perspectives ☆," 2020.

[15] F. Fine, J. L. Lucas, J. M. Chardigny, B. Redlingshöfer, and M. Renard, "Food losses and waste in the French oilcrops sector," *OCL - Oilseeds fats*, 2015, vol. 22, no. 3.

[16] N. Turgut Dunford, *Oxidative Stability of Sunflower Seed Oil*. AOCS Press., 2015.

[17] B. Bonazza, M. Herman, B. Saunders, J. Bennett, E. Kloss, and J. F. da Silva Junior, "White Paper on Internationally Compatible Biofuel Standards - Tripartite Task Force Brazil, European Union & United States of America," pp. 1–95, 2007.

[18] M. Achparaki *et al.*, "We are IntechOpen , the world ' s leading publisher of Open Access books Built by scientists , for scientists TOP 1 %," *Intech*, p. 13, 2012.

[19] M. Othman and R. S. R. Kasim, "Determination of the Effects of Safflower Biodiesel and Its Blends with Diesel Fuel on Engine Performance and Emissions in a Single Cylinder Diesel Engine," *Int. Conf. Softw. Technol. Eng. 3rd (ICSTE 2011)*, vol. 1, pp. 599–607.

[20] S. Fernando, P. Karra, R. Hernandez, and S. K. Jha, "Effect of incompletely converted soybean oil on biodiesel quality," *Energy*, 2007, vol. 32, no. 5, pp. 844–851, 2007.

[21] G. Appiah, S. K. Tulashie, E. E. A. Akpari, E. R. Rene, and D. Dodoo, "Biolubricant production via esterification and transesterification processes: Current updates and perspectives," *Int. J. Energy Res.*2022, vol. 46, no. 4, pp. 3860–3890.

[22] A. Abbaszaadeh, B. Ghobadian, M. R. Omidkhah, and G. Najafi, "Current biodiesel production technologies: A comparative review," *Energy Convers. Manag.*2012, vol. 63, pp. 138–148.

[23] L. M. Diamante and T. Lan, " Absolute Viscosities of Vegetable Oils at Different Temperatures and Shear Rate Range of 64.5 to 4835 s − 1 ," *J. Food Process.*2014, vol. 2014, pp. 1–6.

[24] M. Gülüm and A. Bilgin, "Density, flash point and heating value variations of corn oil biodiesel-diesel fuel blends," *Fuel Process. Technol.*2015, vol. 134, pp. 456–464.

[25] Z. S. Hu *et al.*, "Preparation and tribological properties of nanometer magnesium borate as lubricating oil additive," *Wear*, 2022, vol. 252, no. 5–6, pp. 370–374.

[26] D. X. Peng, Y. Kang, R. M. Hwang, S. S. Shyr, and Y. P. Chang, "Tribological properties of diamond and SiO2 nanoparticles added in paraffin," *Tribol. Int.*2008, vol. 42, no. 6, pp. 911–917.

[27] P. P. Chiplunkar and A. P. Pratap, "Utilization of sunflower acid oil for synthesis of alkyd resin," *Prog. Org. Coatings*,2016, vol. 93, no. April, pp. 61–67.

[28] A. R. Abdulmunem, M. H. Jabal, and H. S. Abd, "IC-engine performance using MWCNT dispersed in non-edible sunflower oil as bio lubricant: Experimental investigation," *Int. J. Mech. Eng. Technol.*2018, vol. 9, no. 13, pp. 1071–1082.

[29] A. Hernández Battez *et al.*, "CuO, ZrO2 and ZnO nanoparticles as antiwear additive in oil lubricants," *Wear*, 2008, vol. 265, no. 3–4, pp. 422–428.

[30] H. M. Mobarak *et al.*, "The prospects of biolubricants as alternatives in automotive applications," *Renew. Sustain. Energy Rev.*2014, vol. 33, no. May, pp. 34–43.

[31] N. J. Fox, B. Tyrer, and G. W. Stachowiak, "Boundary lubrication performance of free fatty acids in sunflower oil," *Tribol. Lett.*2004, vol. 16, no. 4, pp. 275–281.

[32] Y. Singh, A. Farooq, A. Raza, M. A. Mahmood, and S. Jain, "Sustainability of a non-edible vegetable oil based bio-lubricant for automotive applications: A review," *Process Saf. Environ. Prot.* 2017, vol. 111, pp. 701–713.

[33] Z. J. Zhang, D. Simionesie, and C. Schaschke, "Graphite and hybrid nanomaterials as lubricant additives," *Lubricants*, 2014, vol. 2, no. 2, pp. 44–65.

[34] W. K. Shafi, A. Raina, and M. I. Ul Haq, "Friction and wear characteristics of vegetable oils using nanoparticles for sustainable lubrication," *Tribol. - Mater. Surfaces Interfaces*, 2018, vol. 12, no. 1, pp. 27–43.

[35] A. Singh, P. Chauhan, and T. G. Mamatha, "A review on tribological performance of lubricants with nanoparticles additives," *Mater. Today Proc.*2019, vol. 25, no. xxxx, pp. 586–591.

[36] A. D. Thampi, M. A. Prasanth, A. P. Anandu, E. Sneha, B. Sasidharan, and S. Rani, "The effect of nanoparticle additives on the tribological properties of various lubricating oils - Review," *Mater. Today Proc.*2021, vol. 47, no. xxxx, pp. 4919–4924.

[37] Y. Su, L. Gong, and D. Chen, "An investigation on tribological properties and lubrication mechanism of graphite nanoparticles as vegetable based oil additive," *J. Nanomater.*, vol. 2015.

[38] M. A. S. Biswas *et al.*, "Lubrication Performance of Sunflower Oil Reinforced with Halloysite Clay Nanotubes (HNT) as Lubricant Additives," *Lubricants*, 2022, vol. 10, no. 7.

[39] E. Omrani, A. Siddaiah, A. D. Moghadam, U. Garg, P. Rohatgi, and P. L. Menezes, "Ball milled graphene nano additives for enhancing sliding contact in vegetable oil," *Nanomaterials*, 2021, vol. 11, no. 3, pp. 1–24.

[40] S. Baskar, G. Sriram, and S. Arumugam, "Experimental analysis on tribological behavior of nano based bio-lubricants using four ball tribometer," *Tribol. Ind.*2015, vol. 37, no. 4, pp. 449–454.

[41] V. Cortes, K. Sanchez, R. Gonzalez, M. Alcoutlabi, and J. A. Ortega, "The performance of SiO2 and TiO2 nanoparticles as lubricant additives in sunflower oil," *Lubricants*, 2020, vol. 8, no. 1.

INVESTIGATION OF THE EFFECTS OF ETHANOL, ISOPROPANOL, ISOBUTANOL AND DIETHYL ETHER ADDITIVES ADDED TO DIESEL FUEL ON ENGINE VIBRATION AND NOISE IN A SINGLE CYLINDER COMPRESSION IGNITION ENGINE

Murat CINIVIZ[1], Nurullah GULTEKIN[2], Halil Erdi GULCAN[3]

INTRODUCTION

The transportation sector, which has been growing rapidly for the last century, has increased the use of fossil fuels. Countries that want to reduce their dependence on fossil fuels have turned to alternative energy sources. As a result, European countries have increased the use of fuels produced from renewable energy sources [1,2]. The use of these fuels not only reduces the use of fossil fuels, but also provides improvements in performance and emissions. The use of alcohol-type fuels in internal combustion engines directly or in a mixture is also economical as it does not require any modification on the engine. Therefore, studies examining the use of alcohol-diesel mixtures in compression ignition engines have increased in recent years [3,4].

Vibration and noise in the engine affect the comfort of the vehicle as well as the operating performance of the engine [22]. In addition, international standards bring a number of standards to vehicle noise. Since the use of alcohol mixtures in internal combustion engines changes the fuel chemistry, it affects the operation of the engine [8,9,10]. Therefore, alcohol mixtures affect engine performance and emissions, as well as noise and emissions resulting from combustion.

By mixing suitable alcohols with diesel fuel with appropriate mixing ratios and using them in compression ignition engines, the vibration and noise level of the engine can be reduced. Especially reducing the noise and vibrations created by diesel engines will make these vehicles the reason for preference. Because many buyers do not prefer diesel engines because they create higher vibration and noise compared to gasoline engines [23].

Ethanol is the leading alcohol added to diesel fuel. Ethanol is preferred because it contains high oxygen concentration and low sulfur content [5]. Oxygen atoms in ethanol can accelerate combustion by reducing the equivalence ratio in the flame [6]. Taghizadeh et al [7] compared pure diesel fuel with 5% ethanol added diesel

1   Selcuk University, Faculty of Technology, Mechanical Engineering Department Konya/Turkey.
2   Karamanoğlu Mehmetbey University, Vocational School of Technical Sciences, Automotive Technology, Karaman/Turkey.
3   Selcuk University, Faculty of Technology, Mechanical Engineering Department Konya/Turkey.

fuel in their study. In the study, they found that the vibration increased by 4.7% in the experiments with ethanol added fuel. Addition of ethanol to diesel fuel provides a reduction in emissions. However, the low cetane number of ethanol prevents its use in diesel engines [11]. This problem can be avoided by adding diethyl ether fuel with a high cetane number. Another problem of adding ethanol to diesel fuel is solvent. This problem can be solved by adding isopropanol to the fuel. Isopropanol with a long carbon chain can be effective in providing better solubility to the fuel. Liu et al [12] found an increase in engine performance and emissions in their tests with isopropanol added diesel fuel. In another study parallel to this study; Hazar and Uyar [13] carried out a study comparing isopropanol-diesel fuel mixture with pure diesel. Fuels containing different isopropanol additions were used in the study. The addition of isopropanol increases the specific fuel consumption. It has been found to reduce soot emissions.

Isobutanol produced from renewable energy sources can be used by mixing with diesel fuel. In the studies, it has been determined that the addition of isobutanol is effective in reducing the emission values. He [14] investigated the combustion process and performance by adding 10% isobutanol to diesel fuel in his study. In the study, it was determined that isobutanol is quite effective in reducing soot emissions due to its oxygen content. Yang et al [15], who think that isobutanol can help to overcome the environmental and energy crisis, added 10% by volume isobutanol to biodiesel fuel in their study. In the study, they determined that there was a 38.5% reduction in soot emissions due to the biodiesel ratio. Karabektaş and Hoşöz [16] carried out tests by applying a 5-10-15-20% isobutanol mixture to diesel fuel in their experimental study. As a result of the study, it was determined that with the addition of isobutanol, a decrease in CO and NO_X emissions and an increase in HC emissions.

When the studies are evaluated in general, there are many studies on the creation of mixtures of alcohol mixtures with diesel fuel [17,18,19,20,21]. However, there are very few studies examining the effects of alcohol mixtures on the noise and vibration of the engine. It is clear that alcohols will be effective in reducing vibration and noise emissions. Therefore, this is an issue that needs to be worked on.

In this study, tests were carried out with different fuel mixtures in a single-cylinder diesel engine. Tests with alcohol mixtures were compared with standard diesel fuel and the effect of using alcohol mixtures on engine vibration and noise was examined. The main purpose of the study is to reduce vibration and noise values by applying alcohol mixtures to diesel fuel.

MATERIAL AND METHOD

Test Engine

The engine used in the tests is the *ANTOR AD 320* single-cylinder diesel engine shown in **Figure 1**, and its technical specifications are given in **Table 1**.
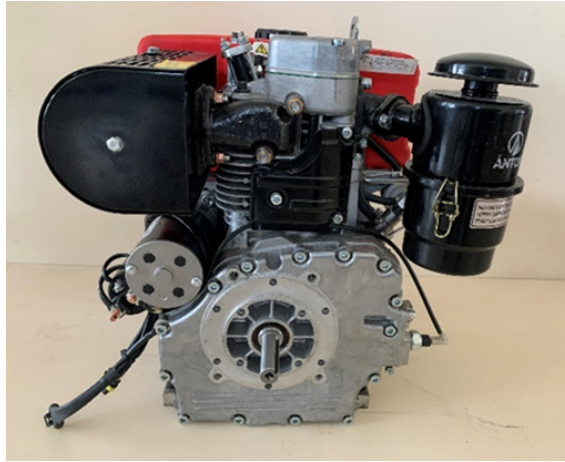
Figure 1. Test engine

Table 1. Engine specifications

Brand/Model	ANTOR / AD 320
Number of cylinders	1
Cylinder volume	315 cm³
Cylinder diameter	78 mm
Stroke	66 mm
Compression ratio	17,5/1
Engine speed	3600 rpm
Max. torque	10,5 @1850 rpm
Crankcase oil capacity	1,2 lt
Injection timing	20 [ºCA bTDC]
Injection pressure	220 [Bar]

Dynamometer

The dynamometer used to measure the torque and power of the engine is the *ABB* brand air-cooled active dynamometer seen in **Figure 2**. Its technical specifications are shown in **Table 2.**
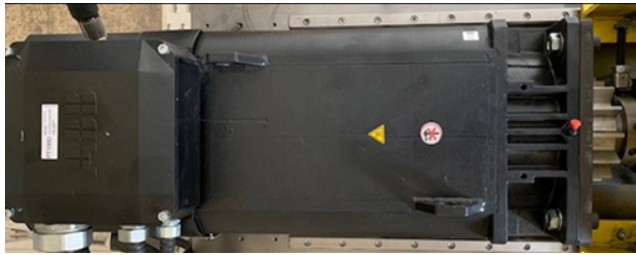
Figure 2. Dynamometer

Table 2. Dynamometer technical specifications

Brand/Model	ABB/Square body
Rated power (kW)	49,3
Rated speed (rpm)	3000
Rated torque (Nm)	157
Maximum speed (rpm)	7500
Power factor	0,831
Yield (%)	93,6

Velocity measurement was made with a 1024 ppr HTL encoder connected to the dynamometer shaft.

Vibration Device

A vibration device was used to measure the oscillation produced by the engine. The *PCE-VD3* model accelerometer vibration device seen in **Figure 3a** was used to record the data. This device is a miniature data logger with an integrated X, Y, Z three-axis acceleration sensor. The built-in sensor in the device can measure within ±18 measuring range per axis. It measures the total acceleration obtained in three (X, Y, Z) axes and in four different acceleration (g) units. All data were automatically recorded at a time interval of 500 ms with the computer interface shown in **Figure 3b**. Data were recorded over a period of 90 seconds.

(a)
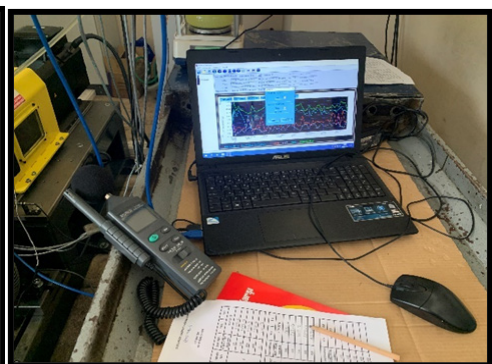
(b)

Figure 3. Vibration device (a) and computer interface (b)

The vibration device is mounted on the cylinder head of the engine. As seen in **Figure 4**, the device is fixed in such a way that it does not move independently. The device is connected to the computer with a USB cable to save the data it receives.

Figure 4. Connecting the vibrator to the motor

Noise Device

The *GERATECH DT 8820* sound level meter shown in **Figure 5** was used to measure the noise generated during engine operation. Values are determined in decibel dB(A). Noise measurements were recorded at a distance of 1.0 meters from the noise center in accordance with the ISO 362-1:2007 standard. [24]. The only source of noise is the engine as the measurements are made indoors.

Figure 5. Noise meter

Test Fuels

The fuels are proportionally mixed. The mixing ratios of the test fuels are given in **Table 3.**

Table 3. The volumetric fractions of the test fuel blends.

Test Fuels	Volumetric fraction (%, v/v)				
	Diesel	Ethanol	Isopropanol	Isobutanol	Diethyl ether
D100	100	-	-	-	-
D85E5İP5İB5	85	5	5	5	-
D82.5E5İP5İB5DEE2.5	82.5	5	5	5	2.5
D80E5İP5İB5DEE5	80	5	5	5	5

The physical and chemical properties of the test fuels are given in **Table 4.**

Table 4. Physical and chemical properties of all test fuels [25,26].

Properties	Diesel	Ethanol	Propanol	Isobutanol	Diethyl ether
Mol. formula	C_nH_m	C_2H_5OH	C_3H_7OH	C_4H_9OH	$C_4H_{10}O$
Mol. wt. (Kg/kmol)	185-212	46.06	60.1	74.1	74,12
Density (kg/m³)	830-840	788	800-805	805-810	713
Cetane num.	45-52	5-15	10-12	-	128
Viscosity at 40ºC (m/s²)	0.0027	0,00012	0.0017	0.0022	0.00023

RESULTS

Valuation of Vibration Data

The average vibration values of the engine as a result of the use of standard diesel and D85E5IP5IB5 fuel formed with a mixture of 85% diesel, 5% ethanol, 5% isopropanol, 5% isobutanol. **Figure 6** and **Figure 7** It can be seen at.
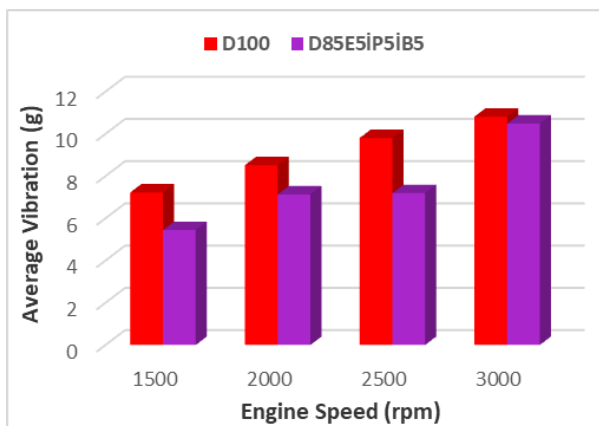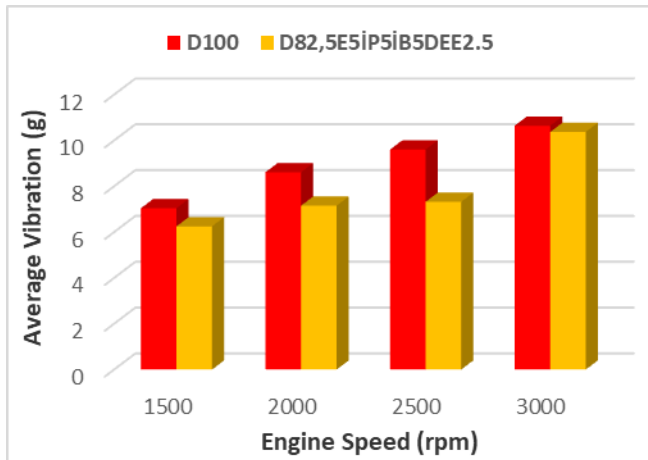

Figure 6. Average vibration generated by D100 and D85E5IP5İB5 fuel at 3 Nm load at different engine speeds.

When the average vibration data determined as a result of the tests performed with standard diesel and D85E5IP5İB5 fuel at 3 Nm load are evaluated.

➢ While the average vibration measured with standard diesel fuel at 1500 rpm is 7.22 g, in the tests performed with D85E5IP5İB5 fuel, the average vibration decreased by 24.7% and was determined as 5.44 g.

➢ While the average vibration measured with standard diesel fuel at 2000 rpm was 8.51 g, in the tests performed with D85E5IP5İB5 fuel, the average vibration decreased by 16.2% and was determined as 7.13 g.

➢ While the average vibration measured with standard diesel fuel at 2500 rpm was 9.8 g, in the tests performed with D85E5IP5İB5 fuel, the average vibration decreased by 26.7% and was determined as 7.19 g.

➢ While the average vibration measured with standard diesel fuel at 3000 rpm is 10.81 g, in the tests performed with D85E5IP5İB5 fuel, the average vibration decreased by 3% and was determined as 10.48 g.
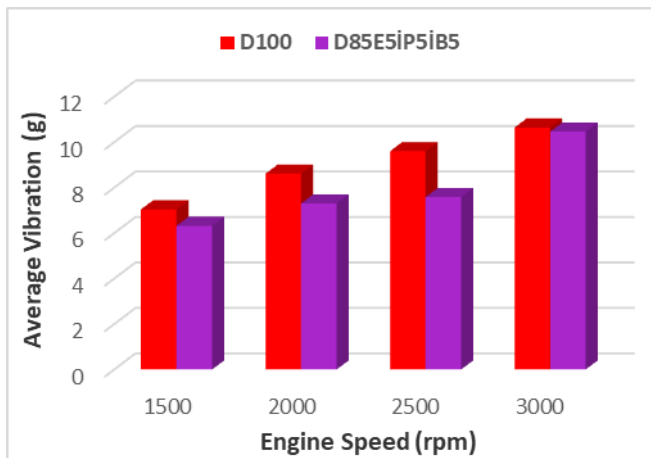
Figure 7. Average vibration generated by D100 and D85E5IP5İB5 fuel at 6 Nm load at different engine speeds.

When the average vibration data determined as a result of the tests performed with standard diesel and D85E5IP5İB5 fuel at 6 Nm load are evaluated.

➢ While the average vibration measured with standard diesel fuel at 1500 rpm is 7.04 g, in the tests performed with D85E5IP5İB5 fuel, the average vibration decreased by 10.2% and was determined as 6.32 g.

➢ While the average vibration measured with standard diesel fuel at 2000 rpm was 8.61 g, in the tests performed with D85E5IP5İB5 fuel, the average vibration was found to be 7.3 g with a decrease of 15.2%.

➢ While the average vibration measured with standard diesel fuel at 2500 rpm was 9.6g, the average vibration was determined to be 7.58g with a decrease of 21% in the tests performed with D85E5IP5İB5 fuel.

➢ While the average vibration measured with standard diesel fuel at 3000 rpm is 10.64 g, in the tests performed with D85E5IP5İB5 fuel, the average vibration decreased by 1.7% and was determined as 10.46 g.

The average vibration values of the engine as a result of the use of D82.5E5IP5İB5DEE2.5 fuel formed with a mixture of standard diesel and 82.5% diesel 5% ethanol, 5% isopropanol, 5% isobutanol, 2.5% diethyl ether **Figure 8** and **Figure 9** It can be seen at.
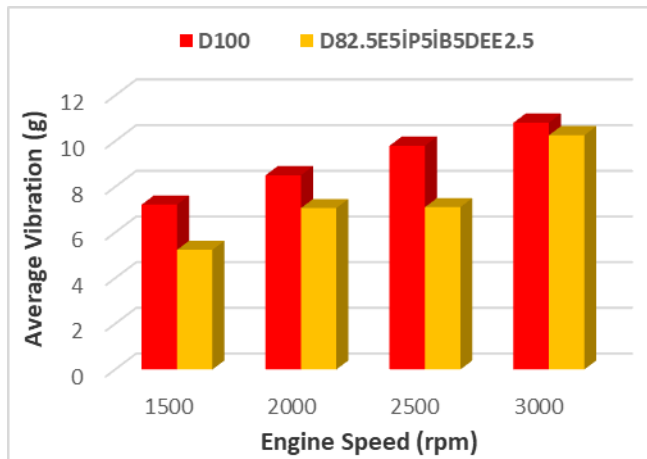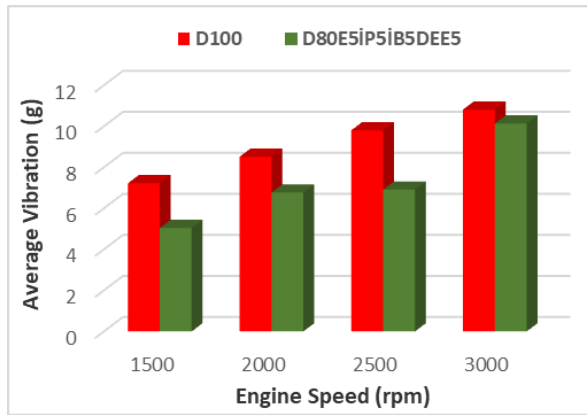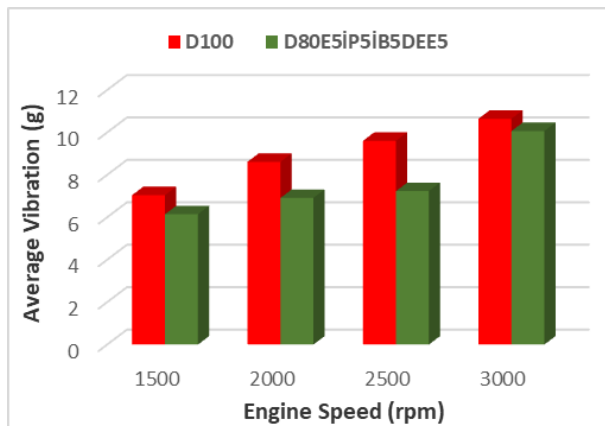

Figure 8. Average vibration generated by D100 and D82.5E5İP5İB5DEE2.5 fuel at 3 Nm load at different engine speeds.

When the average vibration data determined as a result of the tests performed with standard diesel and D82.5E5İP5İB5DEE2.5 fuel at 3 Nm load are evaluated.

➢ While the average was 7.22 g measured with standard diesel fuel at 1500 rpm, it was determined to decrease by an average of 27.4% in the tests performed with D82.5E5IP5İB5DEE2.5 fuel.

➢ While the average was 8.51 g measured with standard diesel fuel at 2000 rpm, it was determined as 7.07 g with an average decrease of 16.9% in the tests performed with D82.5E5IP5İB5DEE2.5 fuel.

➢ While the measurement was 9.8g with standard diesel fuel at 2500 rpm, it was determined as 7.11g with an average decrease of 27.45% in the tests performed with D82.5E5IP5İB5DEE2.5 fuel.

➢ While the average was 10.81g measured with standard diesel fuel at 3000 rpm, it was determined to be 10.26g with an average decrease of 5% in the tests performed with D82.5E5IP5İB5DEE2.5 fuel.

Figure 9. Average vibration generated by D100 and D82.5E5İP5İB5DEE2.5 fuel at 6 Nm load at different engine speeds.

When the average vibration data determined as a result of the tests performed with standard diesel and D82.5E5İP5İB5DEE2.5 fuel at 6 Nm load are evaluated.

- ➢ While the average vibration measured with standard diesel fuel at 1500 rpm is 7.04 g, in the tests performed with D82.5E5IP5İB5DEE2.5 fuel, the average vibration decreased by 11.2% and was determined as 6.25g.

- ➢ While the average vibration measured with standard diesel fuel at 2000 rpm is 8.61 g, in the tests performed with D82.5E5IP5İB5DEE2.5 fuel, the average vibration decreased by 16.96% and was determined as 7.15g.

- ➢ While the average vibration measured with standard diesel fuel at 2500 rpm is 9.6 g, in the tests performed with D82.5E5IP5İB5DEE2.5 fuel, the average vibration was found to be 7.32g with a decrease of 23.75%.

- ➢ While the average vibration measured with standard diesel fuel at 3000 rpm is 10.64 g, in the tests performed with D82.5E5IP5İB5DEE2.5 fuel, the average vibration decreased by 2.5% and was determined as 10.37g.

The average vibration values of the engine as a result of the use of D80E5IP5İB5DEE5 fuel formed with a mixture of standard diesel and 80% diesel, 5% ethanol, 5% isopropanol, 5% isobutanol, 5% diethyl ether. **Figure 10** and **Figure 11** It can be seen at.

Figure 10. Average vibration generated by D100 and D80E5İP5İB5DEE5 fuel at 3 Nm load at different engine speeds.

When the average vibration data determined as a result of the tests performed with standard diesel and D80E5İP5İB5DEE5 fuel at 3 Nm load are evaluated.

➢ While the average vibration measured with standard diesel fuel at 1500 rpm is 7.22 g, in the tests performed with D80E5IP5İB5DEE5 fuel, the average vibration decreased by 30.2% and was determined as 5.04 g.

➢ While the average vibration measured with standard diesel fuel at 2000 rpm is 8.51 g, in the tests performed with D80E5IP5İB5DEE5 fuel, the average vibration decreased by 20.45% and was determined as 6.77g.

➢ While the average vibration measured with standard diesel fuel at 2500 rpm is 9.8 g, in the tests performed with D80E5IP5İB5DEE5 fuel, the average vibration decreased by 29.5% and was determined as 6.91g.

➢ While the average vibration measured with standard diesel fuel at 3000 rpm is 10.81 g, in the tests performed with D80E5IP5İB5DEE5 fuel, the average vibration was determined to be 10.13 g, with a decrease of 6.3%.

Figure 11. Average vibration generated by D100 and D80E5İP5İB5DEE5 fuel at 6 Nm load at different engine speeds.

When the average vibration data determined as a result of the tests performed with standard diesel and D80E5İP5İB5DEE5 fuel at 6 Nm load are evaluated.

➢ While the average vibration measured with standard diesel fuel at 1500 rpm is 7.04 g, in the tests performed with D80E5IP5İB5DEE5 fuel, the average vibration decreased by 12.8% and was determined as 6.14 g.

➢ While the average vibration measured with standard diesel fuel at 2000 rpm is 8.61 g, in the tests performed with D80E5IP5İB5DEE5 fuel, the average vibration decreased by 19.74% and was determined as 6.91g.

➢ While the average vibration measured with standard diesel fuel at 2500 rpm is 9.6 g, in the tests performed with D80E5IP5İB5DEE5 fuel, the average vibration was found to be 7.23g with a decrease of 24.68%.

➢ While the average vibration measured with standard diesel fuel at 3000 rpm is 10.64 g, in the tests performed with D80E5IP5İB5DEE5 fuel, the average vibration decreased by 5.5% and was determined as 10.06 g.

Comparison of the average vibration data of all test fuels **Figure 12** and **Figure 13** It can be seen at.
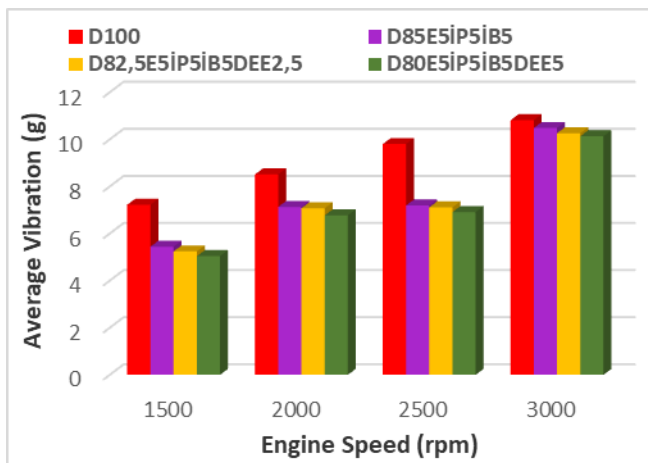
Figure 12. Average vibration of all test fuels at 3 Nm load and different speeds.

Considering the average vibration data of the engine as a result of the use of all test fuels at 3 Nm load;

➢ The lowest vibration at 1500 rpm was determined as 5.04g with D80E5IP5İB5DEE5 fuel.

➢ The lowest vibration at 2000 rpm was determined as 6.77g with D80E5IP5İB5DEE5 fuel.

➢ The lowest vibration at 2500 rpm was determined as 6.91g with D80E5IP5İB5DEE5 fuel.

> ➤ The lowest vibration at 3000 rpm was determined as 10.13g with D80E5IP5İB5DEE5 fuel.
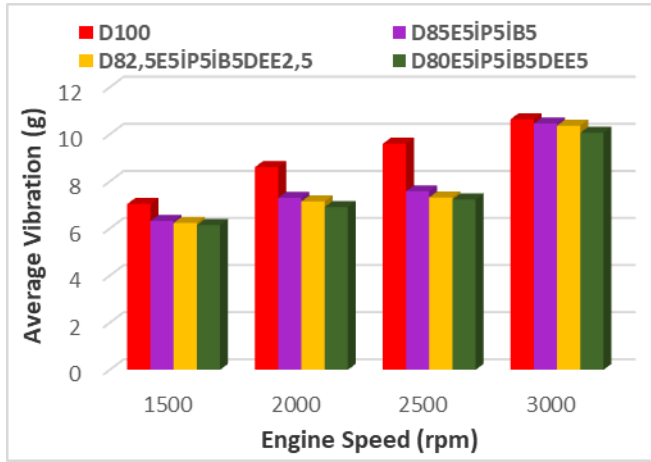
Figure 13. Average vibration of all test fuels at 6 Nm load and different speeds.

Considering the average vibration data of the engine as a result of the use of all test fuels at 6 Nm load;

> ➤ The lowest vibration at 1500 rpm was determined as 6.14 g with D80E5IP5İB5DEE5 fuel.

> ➤ The lowest vibration at 2000 rpm was determined as 6.91g with D80E5IP5İB5DEE5 fuel.

> ➤ The lowest vibration at 2500 rpm was determined as 7.23g with D80E5IP5İB5DEE5 fuel.

> ➤ The lowest vibration at 3000 rpm was determined as 10.06 g with D80E5IP5İB5DEE5 fuel.

When all data are evaluated, it can be concluded that alcohol mixtures are quite effective in reducing vibration values. It was observed that the effect of diethyl ether additive was particularly high. It can be concluded that the high cetane number of diethyl ether reduces the ignition delay and provides a smoother operation.

Evaluation of Noise Data

Noise values generated by the engine as a result of the use of standard diesel and D85E5IP5IB5 fuel, which is formed by a mixture of 85% diesel, 5% ethanol, 5% isopropanol, and 5% isobutanol. **Figure 14** and **Figure 15** It can be seen at.
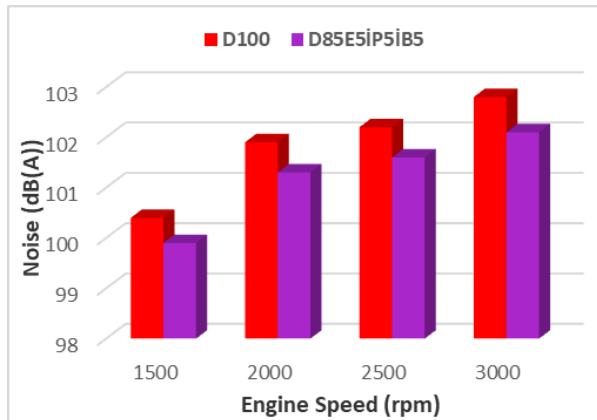
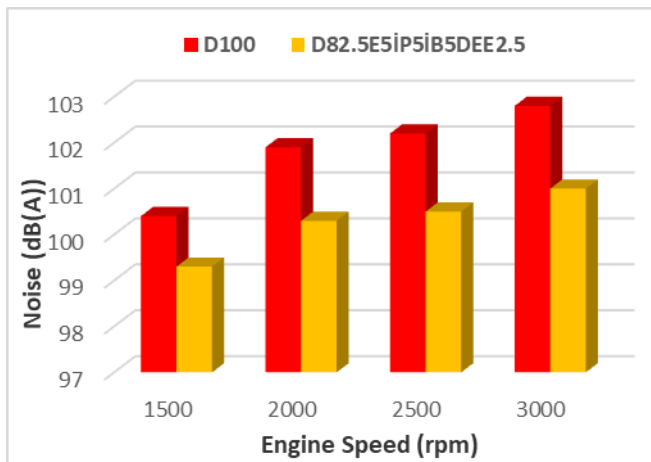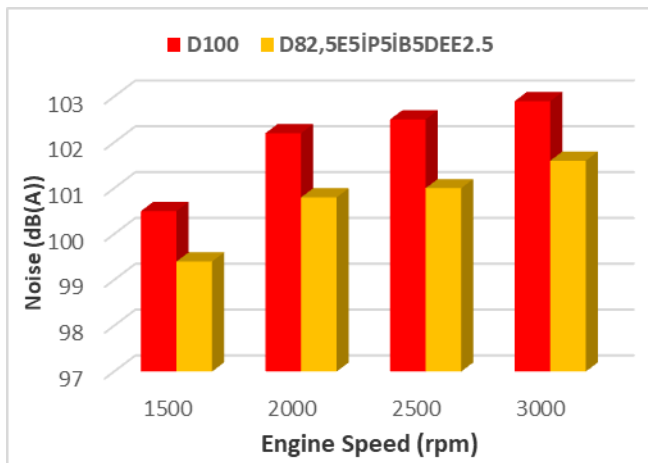**Figure 14.** Noise generated by D100 and D85E5IP5İB5 fuel at 3 Nm load at different engine speeds.

When the noise data determined as a result of the tests performed with standard diesel and D85E5IP5İB5 fuel at 3 Nm load are evaluated;

➢ While the noise value measured with standard diesel fuel at 1500 rpm was 100.4 dBA, in the tests performed with D85E5IP5İB5 fuel, it decreased by 0.5 dBA and became 99.9 dBA.

➢ While the noise value measured with standard diesel fuel at 2000 rpm was 101.9 dBA, in the tests performed with D85E5IP5İB5 fuel, it decreased by 0.6 dBA and became 101.3 dBA.

➢ While the noise value measured with standard diesel fuel at 2500 rpm was 102.2 dBA, in the tests performed with D85E5IP5İB5 fuel, it decreased by 0.6 dBA and became 101.6 dBA.

➢ While the noise value measured with standard diesel fuel at 3000 rpm was 102.8 dBA, in the tests performed with D85E5IP5İB5 fuel, it decreased by 0.7 dBA and became 102.1 dBA.
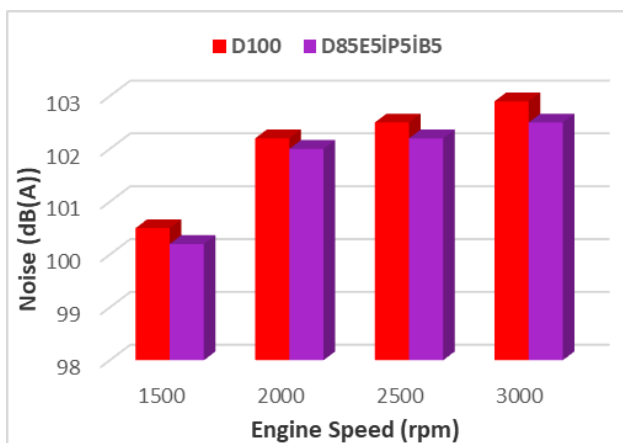
Figure 15. Noise generated by D100 and D85E5IP5İB5 fuel at 6 Nm load at different engine speeds

When the noise data determined as a result of the tests performed with standard diesel and D85E5IP5İB5 fuel at 6 Nm load are evaluated;

➢ While the noise value measured with standard diesel fuel at 1500 rpm was 100.5 dBA, in the tests performed with D85E5IP5İB5 fuel, it decreased by 0.3 dBA and became 100.2 dBA.

➢ While the noise value measured with standard diesel fuel at 2000 rpm was 102.2 dBA, it decreased by 0.2 dBA to 102 dBA in the tests performed with D85E5IP5İB5 fuel.

➢ While the noise value measured with standard diesel fuel at 2500 rpm was 102.5 dBA, in the tests performed with D85E5IP5İB5 fuel, it decreased by 0.3 dBA and became 102.2 dBA.

➢ While the noise value measured with standard diesel fuel at 3000 rpm was 102.9 dBA, in the tests performed with D85E5IP5İB5 fuel, it decreased by 0.4 dBA and became 102.5 dBA.

The noise values generated by the engine as a result of the use of D82.5E5IP5İB5DEE2.5 fuel formed with a mixture of standard diesel and 82.5% diesel 5% ethanol, 5% isopropanol, 5% isobutanol, 2.5% diethyl ether **Figure 16** and **Figure 17** It can be seen at.
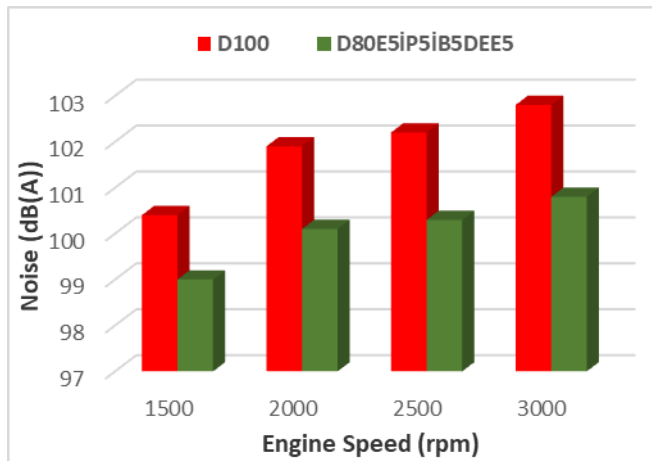

Figure 16. Noise generated by D100 and D82.5E5İP5İB5DEE2.5 fuel at 3 Nm load at different engine speeds.

When the noise data determined as a result of the tests performed with standard diesel and D82.5E5IP5İB5DEE2.5 fuel at 3 Nm load are evaluated;

➢ While the noise value measured with standard diesel fuel at 1500 rpm was 100.4 dBA, it decreased by 1.1 dBA to 99.3 dBA in the tests performed with D82.5E5IP5İB5DEE2.5 fuel.

➢ While the noise value measured with standard diesel fuel at 2000 rpm was 101.9 dBA, it decreased by 1.6 dBA to 100.3 dBA in the tests performed with D82.5E5IP5İB5DEE2.5 fuel.

➢ While the noise value measured with standard diesel fuel at 2500 rpm was 102.2 dBA, it decreased by 1.7 dBA to 100.5 dBA in the tests performed with D82.5E5IP5İB5DEE2.5 fuel.

➢ While the noise value measured with standard diesel fuel at 3000 rpm was 102.8 dBA, it decreased by 1.8 dBA to 101 dBA in the tests performed with D82.5E5IP5İB5DEE2.5 fuel.
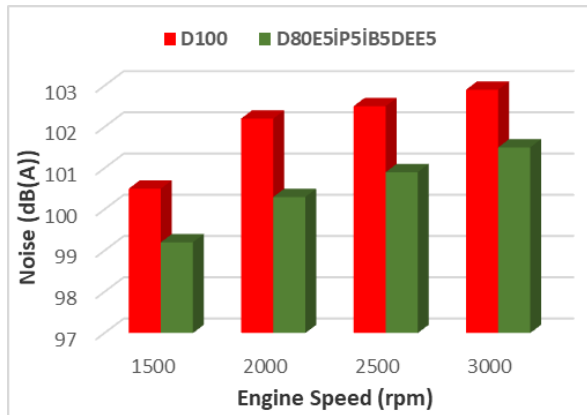

Figure 17. Noise generated by D100 and D82.5E5IP5İB5DEE2.5 fuel at 6 Nm load at different engine speeds.

When the noise data determined as a result of the tests performed with standard diesel and D82.5E5IP5İB5DEE2.5 fuel at 6 Nm load are evaluated;

➢ While the noise value measured with standard diesel fuel at 1500 rpm was 100.5 dBA, it decreased by 1.1 dBA to 99.4 dBA in the tests performed with D82.5E5IP5İB5DEE2.5 fuel.

➢ While the noise value measured with standard diesel fuel at 2000 rpm was 102.2 dBA, in the tests performed with D82.5E5IP5İB5DEE2.5 fuel, it decreased by 1.4 dBA and became 100.8 dBA.

➢ While the noise value measured with standard diesel fuel at 2500 rpm was 102.5 dBA, it decreased by 1.5 dBA to 101 dBA in the tests performed with D82.5E5IP5İB5DEE2.5 fuel.

➢ While the noise value measured with standard diesel fuel at 3000 rpm was 102.9 dBA, it decreased by 1.3 dBA to 101.6 dBA in the tests performed with D82.5E5IP5İB5DEE2.5 fuel.

Noise values generated by the engine as a result of using D80E5IP5İB5DEE5 fuel, which is formed with a mixture of standard diesel and 80% diesel, 5% ethanol, 5% isopropanol, 5% isobutanol, 5% diethyl ether. **Figure 18** and **Figure 19** is seen.

Figure 18. Noise generated by D100 and D80E5İP5İB5DEE5 fuel at 3 Nm load at different engine speeds.

When the noise data determined as a result of the tests performed with standard diesel and D80E5IP5İB5DEE5 fuel at 3 Nm load are evaluated;

- ➢ While the noise value measured with standard diesel fuel at 1500 rpm was 100.4 dBA, it decreased by 1.4 dBA to 99 dBA in the tests performed with D80E5IP5İB5DEE5 fuel.

- ➢ While the noise value measured with standard diesel fuel at 2000 rpm was 101.9 dBA, it decreased by 1.8 dBA to 100.1 dBA in the tests performed with D80E5IP5İB5DEE5 fuel.

- ➢ While the noise value measured with standard diesel fuel at 2500 rpm was 102.2 dBA, it decreased by 1.9 dBA to 100.3 dBA in the tests performed with D80E5IP5İB5DEE5 fuel.

- ➢ While the noise value measured with standard diesel fuel at 3000 rpm was 102.8 dBA, it decreased by 2 dBA to 100.8 dBA in the tests performed with D80E5IP5İB5DEE5 fuel.

Figure 19. Noise generated by D100 and D80E5İP5İB5DEE5 fuel at 6 Nm load at different engine speeds.

When the noise data determined as a result of the tests performed with standard diesel and D80E5IP5İB5DEE5 fuel at 6 Nm load are evaluated;

➢ While the noise value measured with standard diesel fuel at 1500 rpm was 100.5 dBA, it decreased by 1.3 dBA to 99.2 dBA in the tests performed with D80E5IP5İB5DEE5 fuel.

➢ While the noise value measured with standard diesel fuel at 2000 rpm was 102.2 dBA, it decreased by 1.9 dBA to 100.3 dBA in the tests performed with D80E5IP5İB5DEE5 fuel.

➢ While the noise value measured with standard diesel fuel at 2500 rpm was 102.5 dBA, it decreased by 1.6 dBA to 100.9 dBA in the tests performed with D80E5IP5İB5DEE5 fuel.

➢ While the noise value measured with standard diesel fuel at 3000 rpm was 102.9 dBA, it decreased by 1.4 dBA in the tests performed with D80E5IP5İB5DEE5 fuel and became 101.5 dBA.

Comparison of the average vibration data of all test fuels **Figure 20** and **Figure 21** It is seen at.
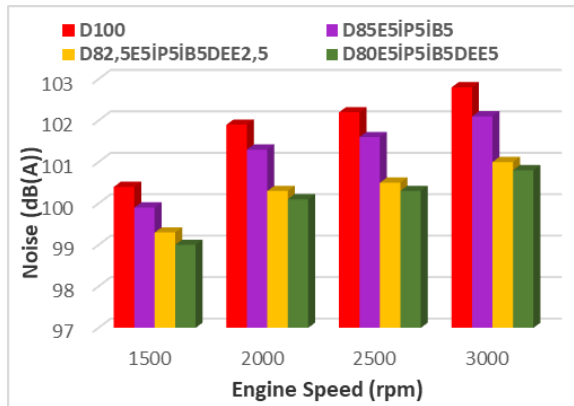
Figure 20. Noise of all test fuels at 3 Nm load and different speeds

When the noise data generated in the engine as a result of the use of all test fuels at a load of 3 Nm are evaluated;

- ➢ The lowest noise at 1500 rpm was determined as 99 dBA with D80E5IP5İB5DEE5 fuel.

- ➢ The lowest noise at 2000 rpm was determined as 100.1 dBA with D80E5IP5İB5DEE5 fuel.

- ➢ The lowest noise at 2500 rpm was determined as 100.3 dBA with D80E5IP5İB5DEE5 fuel.

- ➢ The lowest noise at 3000 rpm was determined as 100.8 dBA with D80E5IP5İB5DEE5 fuel.
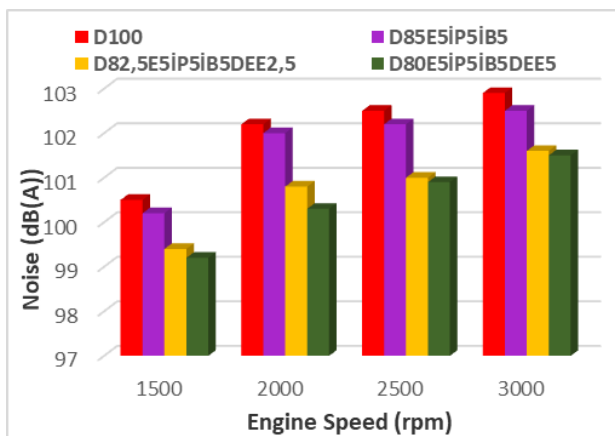
Figure 21. Noise of all test fuels at 6 Nm load and different speeds.

When the noise data generated in the engine as a result of the use of all test fuels at a load of 6 Nm are evaluated;

- ➢ The lowest noise at 1500 rpm was determined as 99.2 dBA with D80E5IP5İB5DEE5 fuel.

➢ The lowest noise at 2000 rpm was determined as 100.3 dBA with D80E5IP5İB5DEE5 fuel.

➢ The lowest noise at 2500 rpm was determined as 100.9 dBA with D80E5IP5İB5DEE5 fuel.

➢ The lowest noise at 3000 rpm was determined as 101.5 dBA with D80E5IP5İB5DEE5 fuel.

When all the data were evaluated, it was determined that alcohol mixtures reduced the noise values. Noise values gave results in the same direction as vibration values. It was observed that the effect of diethyl ether additive was particularly high. It can be concluded that the high cetane number of diethyl ether reduces the ignition delay and provides a smoother operation.

RESULTS AND DISCUSSION

At low and medium loads; It has been determined in the study that diesel mixtures with ethanol, isobutanol and isopropanol added reduce vibration and noise by improving the operation of the engine. However, the low cetane number of alcohols at high loads negatively affects the operation of the diesel engine. Therefore, diethyl ether was added to the mixture to increase the cetane number. The increased number of fuel mixtures improves combustion, thereby reducing noise and vibration at high loads.

➢ In tests performed at 3 Nm load and 2500 rpm, it was determined that D80E5IP5İB5DEE5 fuel reduced engine vibration by 29.5% and engine noise by 1.9 dBA.

➢ In tests performed at 6 Nm load and 2500 rpm, it was determined that D80E5IP5İB5DEE5 fuel reduced engine vibration by 24.7% and engine noise by 1.4 dBA.

➢ The fact that the test engine is a single-cylinder air-cooled diesel engine increases noise and vibration at low and high revs.

➢ The oxygen content of the alcohols used in the mixtures positively affects the combustion at high loads and reduces noise and vibration.

➢ Noise and vibration emissions increase with the increase of engine load and speed. The main reason for this is that more fuel is sent to the cylinder to increase the speed and load. This situation increases the knocking combustion. In order to prevent this, the fuel pressure is increased to ensure that the injected fuel is more atomized. The mechanical fuel system is insufficient in this regard.

ACKNOWLEDGMENT

REFERENCES

[1] Agarwal, AK Biyoyakıtların (alkoller ve biyodizel) içten yanmalı motorlar için yakıt olarak uygulamaları. prog. Enerji Yanması. bilim 2007 , 33 , 233-271. [Google Akademik] [Çapraz Referans]

[2] Armas, Ö.; Cardenas, D.; Garcia-Contreras, R.; Mata, C. Bioethanol-Diesel Blends Used in Diesel Engines and Vehicles Under Temporary Operation. 2020. Available online: https://www. intechopen.com/online-first/bioethanol-diesel-blends-used-in-diesel-engines-and-vehicles-under-transient-operation (date of access 26 March 2021).

[3] Lapuerta, M.; Armas, O.; Garcia-Contreras, R. Stability of diesel-bioethanol blends for use in diesel engines. Fuel 2007, 86, 1351–1357. [Google Scholar] [CrossRef]

[4] Ribeiro, N.M.; Pinto, A.C.; Quintella, C.M.; da Rocha, G.O.; Teixeira, L.S.G.; Guarieiro, L.L.N.; Rangel, M.D.; Veloso, M.C.C.; Rezende, M.J.C.; da Cruz, R.S.; et al. The role of additives for diesel and diesel blended (ethanol or biodiesel) fuels: A review. Energy Fuels 2007, 21, 2433–2445. [Google Scholar] [CrossRef]

[5] Park, S.H.; Youn, I.M.; Lee, C.S. Influence of ethanol blends on the combustion performance and exhaust emission characteristics of a four-cylinder diesel engine at various engine loads and injection timings. Fuel 2011, 90, 748–755. [Google Scholar] [CrossRef]

[6] T.H. Lee, A.C. Hansen, G. Li, T. Lee Effects of isopropanol-butanol-ethanol and diesel fuel blends on combustion characteristics in a constant volume chamber Fuel, 254 (2019), Article 115613

[7] A Taghizadeh-Alisaraei, A Rezaei-Asl. Fuel, 185 (2016).

[8] E. Galloni, F. Scala, G. Fontana Influence of fuel bio-alcohol content on the performance of a turbo-charged, PFI, spark-ignition engine Energy, 170 (2019), pp. 85-92, 10.1016/j. energy.2018.12.129.

[9] D.C. Rakopoulos, C.D. Rakopoulos, D.C. Kyritsis Butanol or DEE blends with either straight vegetable oil or biodiesel excluding fossil fuel: Comparative effects on diesel engine combustion attributes, cyclic variability and regulated emissions trade-off Energy, 115 (2016), pp. 314-325, 10.1016/j.energy.2016.09.022.

[10] H. Wu, Z. Shi, C.-F. Lee, et al. Experimental and kinetic study on ignition of DME/n-butane mixtures under high pressures on a rapid compression machine Fuel, 225 (2018), pp. 35-46, 10.1016/j.fuel.2018.03.129.

[11] K. Owen, T. Coley Automotive fuels reference book (2nd ed.), Warrendale, USA (1995).

[12] Liu Y, Xu B, Jia J, Wu J, Shang W, Ma Z. Effect of injection timing on performance and emissions of di-diesel engine fueled with isopropanol. International Conference on Electrical, Electronics and Mechatronics. 2015;133–137.

[13] H. Hazar, M. Uyar Experimental investigation of isopropyl alcohol (IPA)/diesel blends in a diesel engine for improved exhaust emissions Int J Automotive Eng Tech (IJAET), 4 (2015), pp. 1-6.

[14] Yi Li Yang He. Influence of mixing ratio on performance of isobutanol-diesel fuel [J]. , 2017, 48(10): 41-45.

[15] Po-Ming Yang, Kuang C. Lin, Yuan-Chung Lin, Syu-Ruei Jhang, Shang-Cyuan Chen, Emission evaluation of a diesel engine generator operating with a proportion of isobutanol as a fuel additive in biodiesel blends, Applied Thermal Engineering, Volume 100, 2016, Pages 628-635, ISSN 1359-4311, https://doi.org/10.1016/j.applthermaleng.2016.01.118.

[16] Murat Karabektas, Murat Hosoz,Performance and emission characteristics of a diesel engine using isobutanol–diesel fuel blends,Renewable Energy,Volume 34, Issue 6,2009,Pages 1554-1559,ISSN09601481,https://doi.org/10.1016/j.renene.2008.11.003.

[17] M.R. Chao, T.C. Lin, H.R. Chao, F.H. Chang, C.B. Chen Effects of methanol containing additive on emission characteristics from a heavy-duty diesel engine Sci Total Environ, 279 (2001), pp. 167-179.

[18] N. Yilmaz, T.M. Sanchez Analysis of operating a diesel engine on biodiesel–ethanol and biodiesel–methanol blends Energy, 46 (2012), pp. 126-129.

[19] Z. Guo, T. Li, J. Dong, R. Chen, P. Xue, X. Wei Combustion and emission characteristics of blends of diesel fuel and methanol-to-diesel Fuel, 90 (2011), pp. 1305-1308

[20] C. Yao, C.S. Cheung, C. Cheng, Y. Wang, T.L. Chan, S.C. Lee Effect of diesel/methanol compound combustion on diesel engine combustion and emissions Energy Convers Manage, 49 (2008), pp. 1696-1704.

[21] Gülcan, H. E. , Gültekin, N. & Ciniviz, M. (2022). The Effect of Methanol-Dodecanol Addition on Performance and Smoke Emission in a CI Engine with Diesel Fuel . International Journal of Automotive Science And Technology , 6 (2) , 207-213 . DOI: 10.30939/ijastech..1074513.

[22] Özçelik, Z. & Gültekin, N. (2019). Effect of iridium spark plug gap on emission, noise, vibration of an internal combustion engine . International Journal of Energy Applications and Technologies , 6 (2) , 44-48 . DOI: 10.31593/ijeat.561307.

[23] N. Gültekin , M. Mayda and M. Kilit , " Investigation of the Effect of RPM on Vibration in Gasoline and Diesel Engines", Bitlis Eren University Journal of Science, vol. 6, no. 2, pp. 39-43, Dec. 2017, doi:10.17798/bitlisfen.333042.

[24] https://www.iso.org/standard/42210.html, ISO 362-1:2007 Measurement of noise emitted by accelerating road vehicles — Engineering method — Part 1: M and N categories.

[25] Kumar BR, Saravanan S. Use of higher alcohol biofuels in diesel engines: A review. Renewable and Sustainable Energy Reviews 2016;60:84-115.

[26] Zheng Z, Li C, Liu H, Zhang Y, Zhong X, Yao M. Experimental study on diesel conventional and low temperature combustion by fueling four isomers of butanol. Fuel 2015;141:109-19.

COMPARISON OF JASON-3 TEC WITH GLOBAL IONOSPHERE MAP (GIM)

Gurkan OZTAN[1], Salih ALCAY[2]

INTRODUCTION

Thanks to technological developments, the advancements in satellite systems have gained speed. The increase in the number of analysis centers has brought different techniques and models in data analysis. This increased the accuracy of modeling the Total Electron Content (TEC) in the ionosphere. Global Navigation Satellite System (GNSS), the largest data source for scientific studies on the ionosphere, is important for the evaluation of the changes in the ionosphere. Advances in models and data analysis techniques have made high precision and continuous monitoring at a regional and global scale of the ionosphere [1]–[3]. In monitoring the ionosphere, TEC is produced with the aid of dual-frequency receivers, which are published as global ionosphere maps (GIMs) in the IONosphere EXchange (IONEX) format. GIMs are currently published by eight International GNSS Service (IGS) Ionosphere Associated Analysis Centers (IAACs) [1], [2], [4]–[6]. Besides GIM, TEC data can be derived from different data sources. One of them is radar altimeter satellites. Altimeter satellites have satellite altimeter technique that makes ocean currents, ocean surface winds and waves observation and sea surface measurement globally, frequently and precisely. Dual-frequency altimeters generate an independent source of vertical TEC (VTEC) data [1], [7]. One of the centers providing altimeter data is AVISO. AVISO's first mission, TOPEX/POSEIDON (T/P) is the first major oceanographic survey satellite [3]. The continuation mission of the T/P mission, Jason-1 contains only the fully validated Geophysical Data Record (GDR) [8]. With the termination of Jason-1 mission, OSTM/Jason-2 took over the mission. This new mission introduces two new types of data into the GDR (Geophysical Data Record) data. These data are listed as IGDR (Interim Geophysical Data Record) and OGDR (Operational Geophysical Data Record) [9]. With the expiration of Jason-2, the mission Jason-3 was started in its place. Same the Jason-2, the acquisition of altimeter data continues with the Jason-3 task [10], [11]. Jason provides VTEC directly, independent of GNSS-based VTEC and without the need for any mapping functions [1], [7], [12].

It should be noted that VTEC derived from altimetry data must be filtered for GIM-VTEC comparison. A filter should be used, such as the median filter, which

makes the VTEC generated from the altimeter data comparable to the GIM-VTEC data and eliminates high frequency noise [1].

In this study, VTEC prediction values of Jason-3 altimeter measurements available at https://www.ncei.noaa.gov/data/oceans/ and GIMs in ocean regions where GNSS data is scarce were investigated. Different ionospheric conditions were chosen for the investigation and VTEC values from Jason-3's GDR, OGDR and IGDR data types were used.

DATA AND METHODS

GIM VTEC Data

The ionosphere, which is directly exposed to the sun's rays, is the outermost atmospheric layer of the earth and therefore it is most affected. Free electrons in the ionosphere are of great importance to many disciplines because they affect the signals passing through the atmosphere and have an irregular structure [13]. In order to eliminate or minimize the errors caused by the ionosphere, it makes a great contribution to the precise determination of the number of free electrons in space-based studies. Many organizations such as the IAAC analysis center produce the ionosphere model on a regional and global scale [14]. The dynamic nature of the ionosphere and the modeling of ionospheric parameters such as VTEC need to be continuously improved. The routine production of ionosphere VTEC maps began in 1998 by IAACs IGS. The IONEX format was produced in the same year in order to compare and transform data produced by different centers [15]. Today, there are eight centers producing data and these centers are JPL (Jet Propulsion Laboratory)/U.S.A, CODE (Center for Orbit Determination in Europe)/Switzerland, WHU (Wuhan University)/China, NRCan (Natural Resources Canada)/Canada, ESOC (European Space Operations Center)/Germany, CAS (Chinese Academy of Sciences)/China, UPC (Universitat Politècnica de Catalunya)/Spain, and OPTIMAP (Operational Tool for Ionospheric Mapping And Prediction)/Germany [1], [2], [13], [16]. In this research, data obtained from the CODE/Switzerland data center was used. These centers obtain ionospheric VTEC maps with DCB (differential code deviation) information from GNSS data, along with different ionospheric products [17]. GIM data produced by IAAC centers are used in scientific studies such as space weather and precise positioning. The generated data are obtained by different ionospheric modeling methods. Therefore, they differ in their accuracy and consistency. This indicates that the data need to be verified [1], [5], [18]–[27]. IONEX files are released in different formats. The three types of IONEX files are Quick IONEX files released after twenty-four hours, Forecast IONEX files released in a day or two, and Final IONEX files released in a week or two. Forecast IONEX data is important for minimizing ionospheric errors in applications where real-time and single-frequency receivers are used [14], [28].

IONEX files are available free of charge at https://cddis.nasa.gov/archive/gnss/products/ionex/. The downloaded file directory is YYYY/DDD/AAAAgDDD#.YYi.Z. In this directory, YYYY is the year of the data, DDD is the day of the year, the name of the analysis center is AAA, # is the file number on the relevant day, YY is the last two digits of the relevant year, and .Z is the compressed file [29].

GIM-VTEC data gridded with different approaches and calculation techniques is produced with global models [22]. This can cause accuracy differences between GIMs. Gridding of GIMs is done in different time periods such as 15 minutes, 1 hour and 2 hours for the latitude and longitude to be used. Final IONEX (IGSG) file is published by IGS with an accuracy of 2-8 TECU, taking the average of the GIMs produced by all data centers [1], [13], [14]. In addition to the production of global ionosphere maps, VTEC values obtained from GNSS stations are used in the production of local and regional ionosphere maps. When a study is desired in a certain region, VTEC values of that region can be obtained from these maps. In this study, VTEC values, which are accepted as the basis for comparisons, were used in IONEX format files of the CODE data center. The data can be accessed from the web address http://ftp.aiub.unibe.ch/CODE/. VTEC data can be downloaded in 2-hour resolution until October 2014 and 1-hour resolution for later dates. If the spatial resolution is considered as 2.5° at latitude 5.0° longitude, it has 5183 pieces of data for each hour as ±87.5° latitude and ±180° longitude [13], [15], [29]–[32].

Jason VTEC Data

Information on ocean and sea surfaces is obtained by radar altimeter satellites. TOPEX (TOPography EXperiment)/POSEIDON, AVISO's first mission based in Toulouse, France, is the first major oceanographic survey satellite and was developed by the French space agency CNES (National d'Etudes Spatiale) and the US space agency Center NASA (National Aeronautics and Space Administration) [10]. It features the altimeter technique that measures ocean surface winds, ocean tides, ocean circulation dynamics, geodesy and geodynamics, sea surface measurements and waves in a precise and frequent quadratic scale by altimeter satellites. It also measures the precise height of the sea surface using two state-of-the-art radar altimeter systems. Launched on August 10, 1992, T/P processes and uses level 2 altimeter data. The processed data, including raw data, geophysical data records and sensor data records, are stored in national archives and used in scientific studies. CNES and NASA, which make T/P data available, carry out geophysical arrangements through PO.DAAC and AVISO before opening their archives for scientific studies [3].

The continuation mission of the T/P mission, Jason-1 reworked products launched December 7, 2001, contains only the fully validated GDR. Jason-1 mission replaces previous versions and products. Jason-1 mission terminated on June 21, 2013 [8].

With the termination of Jason-1 mission, OSTM/Jason-2 took over the mission. With the launch of the Jason-2 mission, Meteorological Satellites Exploitation Organization (EUMETSAT) and National Oceanic and Atmospheric Administration (NOAA) agencies were involved in this organization and a more comprehensive altimeter mission was launched. Jason-2, a dual-frequency radar altimeter satellite, is in orbit at an inclination angle of 66° and the satellite's altitude is 1336 km. Provides VTEC data between surface and orbital altitude [33]–[35]. The Jason-2 mission, which ended on October 10, 2019, was launched on June 20, 2008. Jason-2 has the payload of the satellite mission Poseidon-3, the Jason-2 records at 5.3 GHz in the C-band and 13.575 GHz in the Ku band. Information on Jason-2 with a period of 9,9156 days can be obtained at https://www.aviso.altimetry.fr/en/missions/past-missions/jason-2.html [36]. Operational (OGDR) and Interim (IGDR) geophysical data records joined the dataset as OSTM/Jason-2 new products [9].

Similarly, the acquisition of altimeter data is done with Jason-3, the continuation of the OSTM/Jason-2 task. The Jason-3 task was created by the same agencies where the Jason-2 task was carried out, with the Copernicus program created together with the European Union. Jason-3, which was launched on January 17, 2016 and has the same orbit as Jason-2, is 80 seconds apart. Like Jason-2, which has a Poseidon-3 class altimeter, Jason-3 has the same altimeter as Poseidon-3B. After orbiting together for more than six months, Jason-2 and Jason-3 were added to an orbiter between Jason-2's orbit on October 13, 2016. In this case, its features remained the same with its previous orbit, but it caused the ground-track gap to shift by half at equatorial longitudes [35], [37].

The data types of interim and operational and normal geophysical data records, which are level 2 data of the Jason-3 mission, are listed. These products, which are in NetCDF format, contain data about the ocean surface wind speed, ionosphere, important wave height information, troposphere, the ocean with sea surface height and all necessary corrections. It should be noted that the correction data in the file needs filtering. For example, the correction data of the ionosphere should be filtered between -400 millimeters and 40 millimeters. Although the delay amount of the products given in Tab. 1 may differ in type and quantity, the product format is the same [10].

Table 1. The latency and ionospheric correction of Jason-3's operational, intermediate and final geophysical data recording products (O/I/GDR) [10].

Product	OGDR	IGDR	GDR
Delay	3-5 Hour	1-2 Day	~60 Day
Ionosphere Correction	Not available	Available	Available

Jason-2 and Jason-3 move between latitudes approximately 65° south and 65° north. With this movement, it covers almost all ocean regions except high latitude regions. For evaluating GNSS-based VTEC ionospheric models, altimeter data

from Jason satellites would be a very good method. The VTEC data in electrons/m^2 obtained at the Ku-band frequency of the Jason-2 and Jason-3 satellites are derived from the vertical phase ionospheric delay. dR is the Ku band ionospheric gap correction in meters from (O)(I)GDR and the VTEC value is obtained by equation (1) if f_{Ku} is considered as the Ku-band frequency (13.575 GHz) in GHz [38].

$$VTEC = -\frac{dR * f_{Ku}^2}{40.3} \tag{1}$$

RESULTS AND ANALYSIS

In this study, VTEC values of CODE GIMs and Jason-3 (I)(O)GDR were examined with a comparative approach. Jason data, which contains information on ocean regions, provides VTEC data for these regions. VTEC data were obtained from two different data sources for ocean regions on solar active (04/09/2017), geomagnetic active (08/09/2017) and calm (03/04/2022) days. The daily ground tracks of the Jason-3 satellite for the experimental days are given in Fig.1.
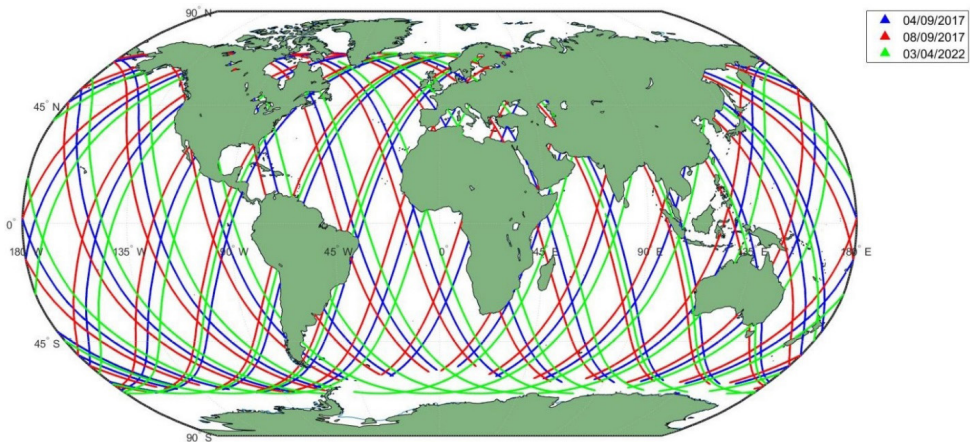
Figure 1. Daily ground tracks of Jason-3 satellite on 04/09/2017, 08/09/2017 and 03/04/2022

The Kp, Dst, and F10.7 indices are measures that reflect the magnitude of changes in the ionosphere. In this study, three different days were selected according to activity levels, taking into account the threshold values of the indices (Fig. 2, Fig. 3 and Fig. 4).

While the F10.7 index value is 185 on the solar active day, the Kp and Dst indices are below the threshold values (Fig. 2). While the Kp value reaches 8 and the Dst index falls below -100 on the geomagnetic active day (Fig. 3), all indices are below the threshold values on the calm day (Fig. 4).
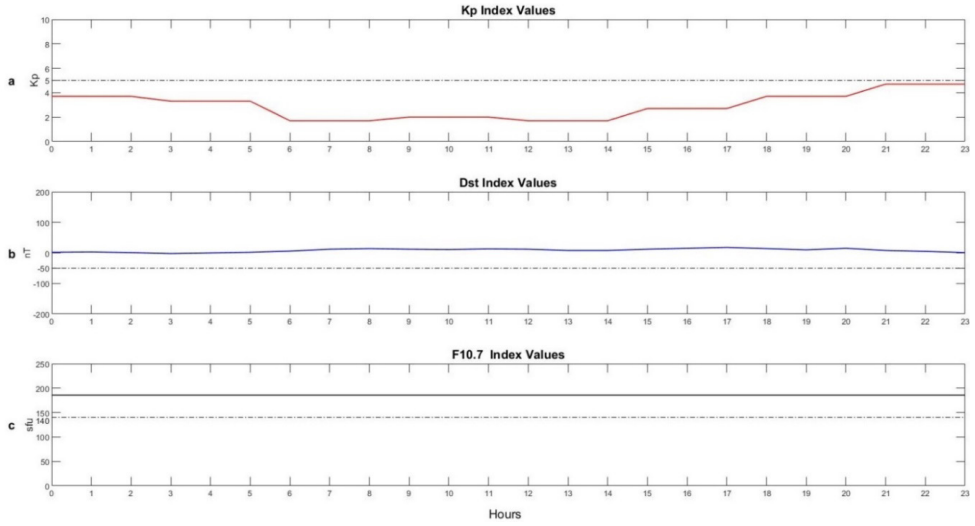
Figure 2. Kp (a) Dst (b) and F10.7 (c) indices on solar active day (04/09/2017).
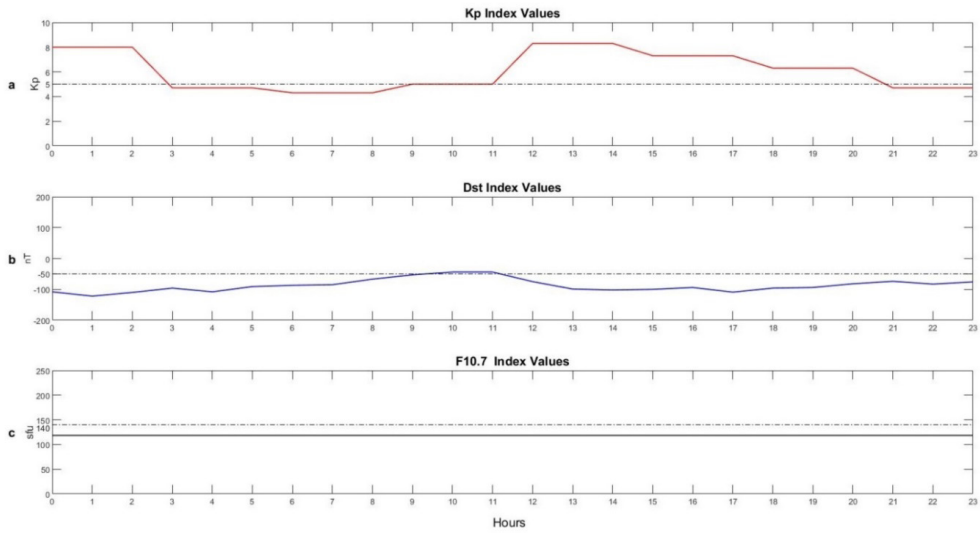
Figure 3. Kp (a), Dst (b) and F10.7 indices (c) on geomagnetic active day (08/09/2017)
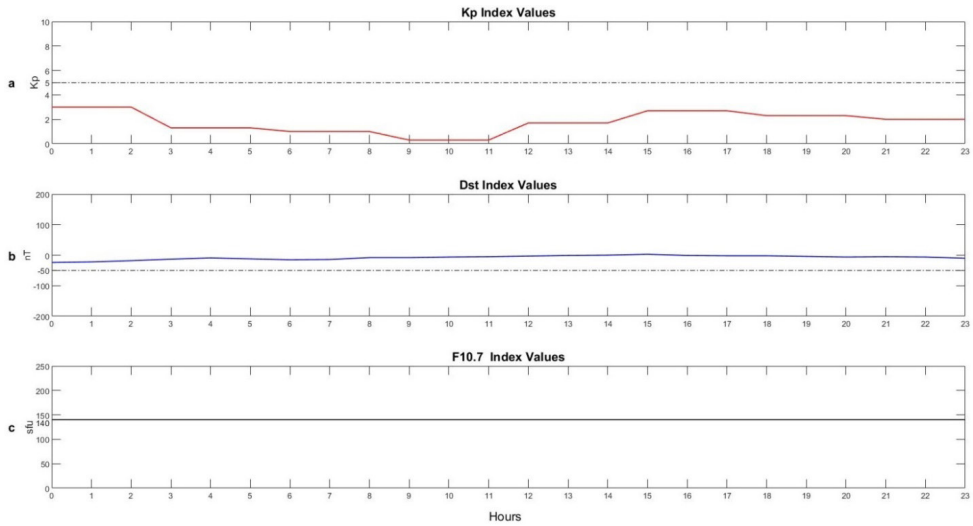
Figure 4. Kp (a), Dst (b) and F10.7 (c) indices on calm day (03/04/2022)

The VTEC values obtained from three different data types GDR, IGDR and OGDR in Jason-3 and VTEC values from GIMs are given in Fig. 5, Fig. 6 and Fig. 7 for solar active, geomagnetic active and calm days, respectively. Although the Jason-3 GDR, IGDR and OGDR VTEC values show similar trends on the calm day (Fig. 7), outliers are observed in IGDR and OGDR VTEC values on active days (Fig. 5 and Fig. 6). In order to examine the magnitude of GIM VTEC and Jason-3 (I)(O)GDR VTEC differences, basic statistical values such as maximum, minimum, mean, and RMS values were computed and given in Tab. 2. As seen in Tab. 2, when the maximum values are examined on the geomagnetic active day, the highest difference belongs to the IGDR and OGDR values. The differences for GDR, IGDR, and OGDR seem to be of similar magnitude during the calm and solar active days. When the mean of the differences is examined, it is observed that the GIM-VTEC values are more compatible with the Jason-3 GDR VTEC values. The mean of the differences ranges between 3.48-4.19, 3.57-6.44, and 4.33-5.39 for GDR, IGDR, and OGDR, respectively. In addition, the RMS values for the three options are below 3 TECU on the calm day.

Solar Active Day 04.09.2017 JASON-VTEC and GIM-VTEC
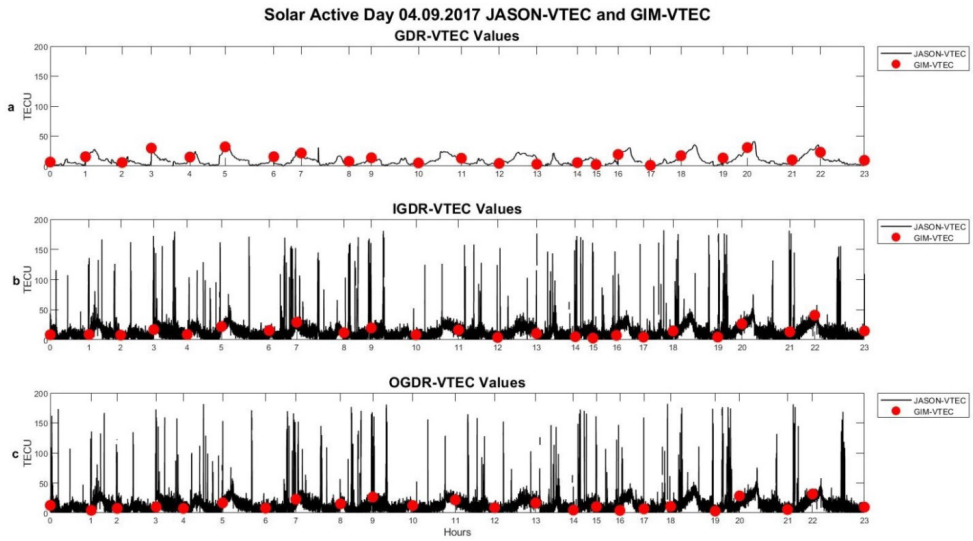
Figure 5. Jason-VTEC GDR (a), IGDR (b), OGDR (c) values and GIM VTEC values on 04/09/2017

Geomagnetic Active Day 08.09.2017 JASON-VTEC and GIM-VTEC
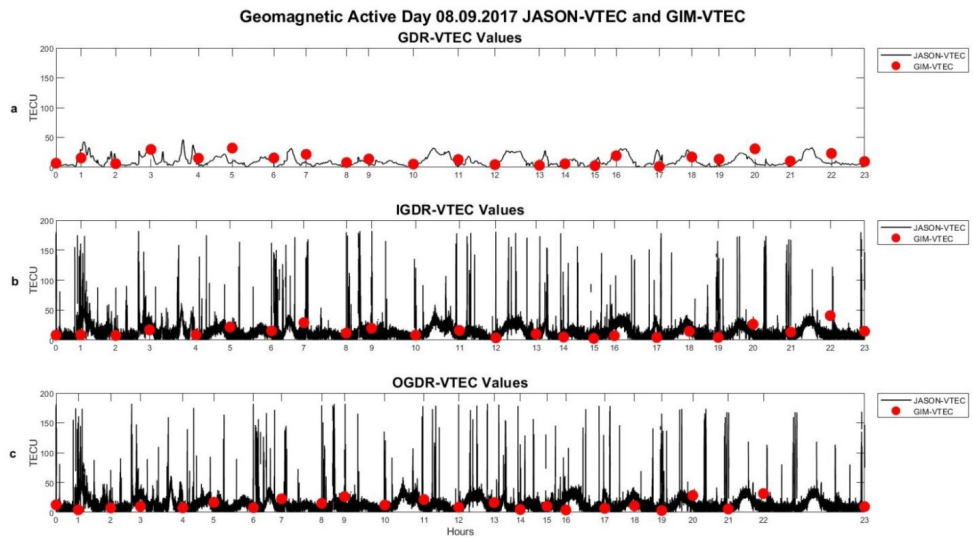
Figure 6. Jason-VTEC GDR (a), IGDR (b), OGDR (c) values and GIM VTEC values on 08/09/2017
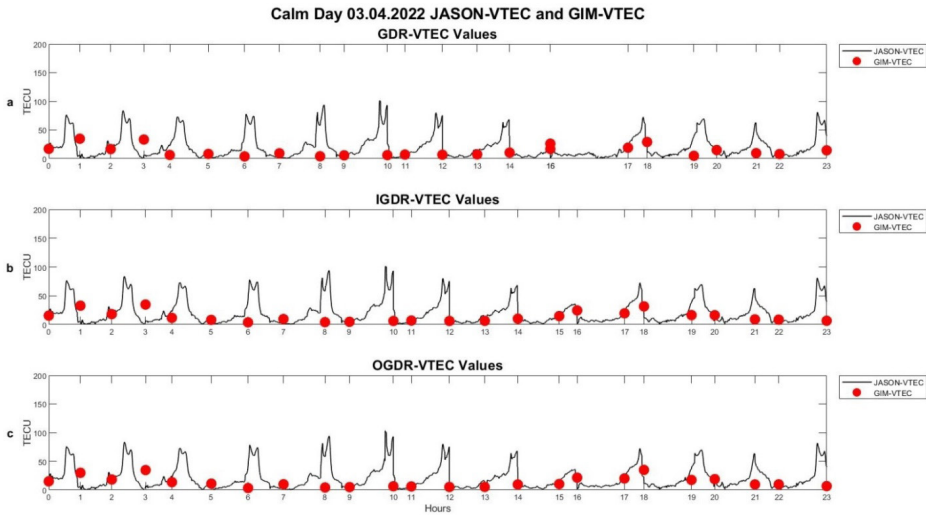
Figure 7. Jason-VTEC GDR (a), IGDR (b), OGDR (c) values and GIM VTEC values on 03/04/2022

Table 2. Basic statistical values of GIM VTEC and Jason VTEC differences

	GIM- JASON								
	4.09.2017			8.09.2017			3.04.2022		
	GDR	IGDR	OGDR	GDR	IGDR	OGDR	GDR	IGDR	OGDR
Maximum	10.01	11.39	12.82	10.74	25.07	19.54	8.94	9.29	10.63
Minimum	0.06	0.03	0.08	0.03	0.10	0.36	0.47	0.47	0.76
Mean	3.48	3.57	4.33	3.74	6.44	5.39	4.19	4.70	5.21
RMS	2.23	2.62	3.51	2.87	5.53	4.58	2.26	2.59	2.86

CONCLUSIONS

In this study, the magnitude of VTEC differences between two different data sources (GIM and Jason-3 (GDR/IGDR/OGDR)) was examined. Since Jason data is produced only for ocean regions, GIM-VTEC and Jason-3 VTEC comparisons were made in these regions on active and calm days. According to the results, Jason-3 GDR VTEC values are more compatible with GIM VTEC values on active and calm days. The mean of the differences between GIM VTEC and the three options of Jason-3 VTEC are mostly below 5 TECU. The results also showed that the magnitude of the differences increases on active days (geomagnetic/solar) compared to the calm day.

ACKNOWLEDGMENTS

REFERENCES

[1]. B. Milanowska, P. Wielgosz, A. Krypiak-Gregorczyk, and W. Jarmołowski, Accuracy of Global Ionosphere Maps in Relation to Their Time Interval, Remote Sens..2021, vol. 13, no. 18, p. 3552, doi: 10.3390/rs13183552.

[2]. Ionosphere – International GNSS Service. 2022, Accessed Oct. 22, 2022, Available: https://igs. org/wg/ionosphere/#charter.

[3]. CLS. and CNES., AVISO USER HANDBOOK MERGED TOPEX/POSEIDON PRODUCTS (GDR-Ms). Jul. 1996.

[4]. G. Johnston, A. Riddell, and G. Hausler, The International GNSS Service, Springer Handbooks. 2017, pp. 967–982, doi: 10.1007/978-3-319-42928-1_33/COVER.

[5]. H. Liu, D. Mei, G. Xu, P. Yang, X. Ren, and X. Zhang, Evaluation and validation of various rapid GNSS global ionospheric maps over one solar cycle, Adv. Sp. Res. Oct.. 2022, vol. 70, no. 8, pp. 2494–2505, doi: 10.1016/J.ASR.2022.07.008.

[6]. P. Wielgosz, B. Milanowska, A. Krypiak-Gregorczyk, and W. Jarmołowski, Validation of GNSS-derived global ionosphere maps for different solar activity levels: case studies for years 2014 and 2018, GPS Solut.. , Jul. 2021, vol. 25, no. 3, pp. 1–15, doi: 10.1007/S10291-021-01142-X/FIGURES/15.

[7]. M. Hernández-Pajares, D. Roma-Dollase, A. Krankowski, A. García-Rigo, and R. Orús-Pérez, Methodology and consistency of slant and vertical assessments for ionospheric electron content models, J. Geod..Dec. 2017, vol. 91, no. 12, pp. 1405–1414, doi: 10.1007/S00190-017-1032-Z/FIGURES/10.

[8]. CNES and JPL, Jason-1 Products Handbook. Apr. 2016, Accessed: Oct. 23, 2022, Available: https://www.aviso.altimetry.fr/fileadmin/documents/data/tools/hdbk_j1_gdr .pdf.

[9]. CNES, EUMETSAT, JPL, and NOAA/NESDIS, OSTM/Jason-2 Products Handbook. Dec. 2011.

[10]. NASA and CNES, Jason-3 Products Handbook. 2018, Available: https://www. ospo.noaa.gov/Products/documents/hdbk_j3.pdf.

[11]. S. Karimi, M. A. Sharifi, S. Farzaneh, and M. Kosary, Combination of swarm, Jason-3, and GNSS observations to construct a new modeling of global ionospheric maps, J. Atmos. Solar-Terrestrial Phys..Nov. 2022, vol. 238–239, p. 105934, doi: 10.1016/J.JASTP.2022.105934.

[12]. G. Jee, H. B. Lee, Y. H. Kim, J. K. Chung, and J. Cho, Assessment of GPS global ionosphere maps (GIM) by comparison between CODE GIM and TOPEX/Jason TEC data: Ionospheric perspective, J. Geophys. Res. Sp. Phys.. Oct. 2010, vol. 115, no. A10, p. 10319, doi: 10.1029/2010JA015432.

[13]. G. Öztan, GPS-TEC değerleri ile IRI-2012 ve IRI-PLAS modellerinin TEC kestirimlerinin karşılaştırılması, Yüksek Lisans Tezi. 2018, pp. 1–142.

[14]. E. Şentürk, Küresel İyonosfer Haritalarının Türkiye'deki Performansının GNSS Verileriyle İstatistiksel Olarak İncelenmesi, Deu Muhendis. Fak. Fen ve Muhendis.. Jan. 2021, vol. 23, no. 67, pp. 247–255, doi: 10.21205/deufmd.2021236721.

[15]. S. Schaer, W. Gurtner, and J. Feltens, IONEX: The IONosphere Map EXchange Format Version 1.1. 1998.

[16]. A. Krypiak-Gregorczyk et al., Validation of DGFI-TUM's new ionosphere model: case studies for year 2018, EGU Gen. Assem. 2022, doi: 10.1007/S10291-021-01142-X.

[17]. J. Feltens and S. Schaer, IGS Products for the Ionosphere, IGS Analysis Centers Workshop ESOC, Darmstadt, Germany. 1998, pp. 1–7.

[18]. F. Azpilicueta and B. Nava, A different view of the ionospheric winter anomaly, Adv. Sp. Res.. Jan. 2021, vol. 67, no. 1, pp. 150–162, doi: 10.1016/J.ASR.2020.10.039.

[19]. K. G. Ratovsky, M. V. Klimenko, Y. V. Yasyukevich, V. V. Klimenko, and A. M. Vesnin, Statistical Analysis and Interpretation of High-, Mid- and Low-Latitude Responses in Regional Electron Content to Geomagnetic Storms, Atmos. 2020, Vol. 11, Page 1308, vol. 11, no. 12, p. 1308, doi: 10.3390/ATMOS11121308.

[20]. J. Zhao, M. Hernández-Pajares, Z. Li, N. Wang, and H. Yuan, Integrity investigation of global ionospheric TEC maps for high-precision positioning, J. Geod.. Mar. 2021, vol. 95, no. 3, pp. 1–15, doi: 10.1007/S00190-021-01487-8/FIGURES/8.

[21]. A. Ciećko and G. Grunwald, Klobuchar, NeQuick G, and EGNOS Ionospheric Models for GPS/ EGNOS Single-Frequency Positioning under 6–12 September 2017 Space Weather Events, Appl. Sci.. 2020, Vol. 10, Page 1553, vol. 10, no. 5, p. 1553, doi: 10.3390/APP10051553.

[22]. M. Hernández-Pajares et al., The IGS VTEC maps: A reliable source of ionospheric information since 1998, J. Geod.. 2009, vol. 83, no. 3–4, pp. 263–275,doi: 10.1007/S00190-008-0266-1.

[23]. D. Bilitza and B. Reinisch, Preface: International reference ionosphere and global navigation satellite systems, Adv. Sp. Res.. Apr. 2015, vol. 55, no. 8, p. 1913, doi: 10.1016/J.ASR.2015.02.021.

[24]. Z. S. Li, N. B. Wang, M. Li, K. Zhou, Y. Bin Yuan, and H. Yuan, Evaluation and analysis of the global ionospheric TEC map in the frame of international GNSS services, Acta Geophys. Sin.. Oct. 2017, vol. 60, no. 10, pp. 3718–3729, doi: 10.6038/CJG20171003.

[25]. D. Roma-Dollase et al., Consistency of seven different GNSS global ionospheric mapping techniques during one solar cycle, J. Geod.. Jun. 2018, vol. 92, no. 6, pp. 691–706, doi: 10.1007/ S00190-017-1088-9/TABLES/4.

[26]. P. Chen, Y. Yao, and W. Yao, Global ionosphere maps based on GNSS, satellite altimetry, radio occultation and DORIS, GPS Solut.. Apr. 2017, vol. 21, no. 2, pp. 639–650, doi: 10.1007/ S10291-016-0554-9/FIGURES/15.

[27]. N. Wang, Z. Li, Y. Yuan, and X. Huo, BeiDou Global Ionospheric delay correction Model (BDGIM): performance analysis during different levels of solar conditions, GPS Solut.. Jul. 2021, vol. 25, no. 3, pp. 1–13, doi: 10.1007/S10291-021-01125-Y/FIGURES/13.

[28]. M. Li, Y. Yuan, N. Wang, Z. Li, and X. Huo, Performance of various predicted GNSS global ionospheric maps relative to GPS and JASON TEC data, GPS Solut. Apr. 2018, vol. 22, no. 2, pp. 1–11, doi: 10.1007/S10291-018-0721-2/TABLES/7.

[29]. CDDIS: NASA's Archive of Space Geodesy Data, NASA. 2022, Accessed May 22, Available: https://cddis.nasa.gov/archive/gnss/products/ionex/.

[30]. Index of /CODE/. 2022, Accessed May 24, 2022, Available: http://ftp.aiub.unibe.ch/ CODE/

[31]. N. Arslan, GPS İle İyonosfer Toplam Elektron Yoğunluğu Değişimlerinin Koordinatlara Etkilerinin Araştırılması. 2004.

[32]. S. Alçay and G. Öztan, Manyetik Fırtınalı ve Fırtınasız Günlerde IRI-PLAS ve IRI-2012 Modellerinin TEC Kestirim Performanslarının İncelenmesi, Harit. Teknol. Elektron. Derg.. 2016, vol. 8, no. 2, pp. 131–140, doi: 10.15659/hartek.16.08.333.

[33]. D. A. Imel, Evaluation of the TOPEX/POSEIDON dual-frequency ionosphere correction, J. Geophys. Res. Ocean. Dec.. 1994, vol. 99, no. C12, pp. 24895–24906, doi: 10.1029/94JC01869.

[34]. E. Astafyeva, I. Zakharenkova, and M. Förster, Ionospheric response to the 2015 St. Patrick's Day storm: A global multi-instrumental overview, J. Geophys. Res. Sp. Phys.. Oct. 2015, vol. 120, no. 10, pp. 9023–9037, doi: 10.1002/2015JA021629.

[35]. S. Biancamaria et al., Validation of Jason-3 tracking modes over French rivers, Remote Sens. Environ.. May 2018, vol. 209, pp. 77–89, doi: 10.1016/J.RSE.2018.02.037.

[36]. Jason-2: Aviso+, Aviso. 2022, Accessed May 23, 2022, Avaliable: https://www.aviso.altimetry.fr/ en/missions/past-missions/jason-2.html

[37]. M. Geodesy et al., Poseidon-3 Radar Altimeter: New Modes and In-Flight Performances,. 2010, vol. 33, no. S1, pp. 53–79, doi: 10.1080/01490419.2010.488970.

[38]. X. Ren, J. Chen, X. Li, X. Zhang, and M. Freeshah, Performance evaluation of real-time global ionospheric maps provided by different IGS analysis centers, GPS Solut.. 2019, vol. 23, no. 4, pp. 1–17, doi: 10.1007/s10291-019-0904-5.

EPOCH AVAILABILITY ANALYSIS OF LOW-COST GNSS RECEIVER

Ceren Konukseven[1], Salih Alcay[2], Sermet Ogutcu[3], Behlül Numan Ozdemir[4],
Mehmet Hacıbeyoglu[5], Esma Nisa Candan[6]

INTRODUCTION

The Global Positioning System (GPS) is the first satellite positioning system that became fully available for public use in 2000 [1]. Following GPS, GLONASS in 2010, Galileo in 2018, and BEIDOU in 2020 reached full operational capacity (FOC). Thanks to the positioning accuracy it provides, the GNSS technique is used in many engineering applications and varied fields. Nevertheless, its application may be limited due to high-priced instruments. Thus, several GNSS manufacturers started to offer low-cost devices. These devices, which were first produced as single-frequency, have recently started to be produced as dual-frequency. Compared with expensive geodetic receivers, low-cost receivers have many advantages in terms of small size, low power consumption, portability, etc. [2].

Numerous studies have been conducted to test the performance of the low-cost GNSS receiver in several ways. Polizzi et al. [3] have developed a monitoring system with several low-cost GNSS receivers and they have tested it in the field. They examined the system in terms of slow and early displacements. As a result of the tests, they have found that cheap GNSS receivers can be used, given the high cost of geodetic receivers. Knight et al. [4] have developed a low-cost GNSS platform for coastal sea level measurement. In their comparisons with more expensive GNSS receivers, they found that similar accuracy performance was achieved. Manzini et al. [1] conducted a study on the suspended bridge to test low-cost GNSS receivers for structural health monitoring applications. They tested the data using RTKlib. They have proven that 1 Hz GNSS solutions with appropriate process parameters can reveal sub-centimetric displacements and 1 cm oscillations at frequencies up to 0.25 Hz. As a result, they have confirmed that low-cost receivers provide data with sufficient accuracy to monitor structures such as bridges and high-rise buildings. Garrido-Carretero et al. [5] evaluated the RTK performance of a dual-frequency geodetic receiver and a single-frequency low-cost GNSS receiver using the ISO-17123-8 standard. Although the geodetic receiver results are better than the low-cost

1   Ceren Konukseven, Necmettin Erbakan University, 0000-0001-6598-9479
2   Salih Alcay, Assoc. Prof. Dr., Necmettin Erbakan University, 0000-0001-5669-7247
3   Sermet Ogutcu, Assoc. Prof. Dr., Necmettin Erbakan University, 0000-0002-2680-1856
4   Behlül Numan Ozdemir, Res. Assist. Dr., Konya Technical University, 0000-0001-7351-1870
5   Mehmet Hacıbeyoglu, Assoc. Prof. Dr., Necmettin Erbakan University, 0000-0003-1830-8516
6   Esma Nisa Candan, Necmettin Erbakan University, 0000-0002-9746-3495

GNSS receiver, it has been found that the low-cost GNSS receiver gives comparable positioning performance. Hamza et al. [6] conducted research utilizing low-cost GNSS receivers and relative and absolute positioning methods to determine the size of the displacements and determine the sensitivity of the three distinct antennas they chose. Additionally, they observationally compared the performance of geodetic and low-cost GNSS receivers. They have determined that a low-cost GNSS receiver can be employed even though the geodetic receiver performs better. Tunini et al. [7] tested a geodetic receiver and a dual-frequency low-cost receiver connected to the same geodetic antenna to monitor crustal deformations. They chose north-eastern Italy as the test site and processed the data with the observations they obtained from the permanent GNSS station in the region. They analyzed the results with time series and obtained that the results of the two receivers are comparable and that the low-cost GNSS receiver provides mm precision. Wielgocka et al. [8] evaluated the signal acquisition performance and accuracy of a dual frequency low-cost GNSS receiver for RTK, NRTK, and static positioning in land survey applications such as the establishment of control points. They concluded that the positioning accuracy obtained by low-cost dual-frequency receivers in static and kinematic modes is sufficient for field surveying applications that require sub-decimeter horizontal accuracy. Romero et al. [9] compared the positioning accuracy obtained from precise point positioning (PPP) and static-relative methods in urban areas using low-cost GNSS receivers. They found that the static-relative method provides a better solution for urban areas. They concluded that the PPP method is suitable for open environments.

Low-cost GNSS receivers have also been employed in other fields, including static-kinematic positioning [10], precise water level measurements [11], RT-PPP application [12], performance analysis in different ionospheric disturbance periods [13], short-long baseline RTK applications [14], monitoring of slope instabilities [15].

In this study, the performance of the dual-frequency low-cost U-blox F9P GNSS receiver was tested in terms of the epoch availability of the RINEX observation file.

MATERIALS AND METHODS

In order to examine the epoch availability of a low-cost GNSS receiver, dual-frequency uBlox F9P was used. Low cost GNSS receiver equipment external antenna is given in Fig. 1 and patch antenna is given in Fig. 2. In this study, the external antenna was used and its properties are provided in Fig.3. In Fig. 3. shows the Environmental data, quality, antenna and mechanical information of the Survey multiband GNSS antenna.

In Fig. 4, a low-cost dual-frequency GNSS receiver (uBlox ZED F9P) is given. The features of the ZED-F9P module such as receiver type and frequency, acquisition, sensitivity, supported protocol types and product size are given in Fig.

5. Low-cost and geodetic receivers were placed on an apparatus close to each other on the roof of Necmettin Erbakan University, Engineering Faculty (Fig. 6).

Figure 1. External antenna

Figure 2. Patch antenna

Antenna	Frequency	GPS L1/L2
		GLONASS L1/L2
		BeiDou B1/B2/B3
		Galileo E1/E2/E5b/E6
	Polarization	RHCP
	Axial ratio	≤3dB
	V.S.W.R.	≤1.5
	Peak gain	≥5dBi
	Impedance	50Ω
	Phase center error	±3mm
	Horizontal coverage angle	360°
LNA	Gain	40±2dB
	Noise figure	≤1.5dB
	Passband fluctuation	±1dB
	Supply voltage	3~5.5V DC
	Current consumption	≤45mA
	V.S.W.R.	≤2.0
Mechanical	Connector	TNC
	Radome material	ABS
	Mounting method	Screw
Environmental	Operating temperature	-40°C~+85°C
	Relative humidity	Up to 95%
	Ingress protection	IP66 (exclude Air vent in screw hole)
	Environmentally friendly	ROHS Compliant

Figure 3. Survey multiband GNSS antenna features [16]
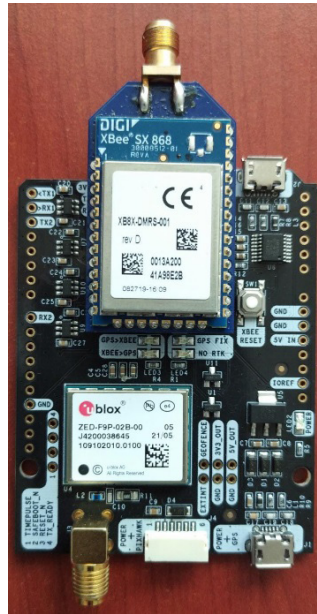
Figure 4. Low-cost dual-frequency GNSS receiver (uBlox ZED-F9P)

	184-channel u-blox F9 engine	
	GPS	L1C/A L2C
	GLONASS	L1OF L2OF
Receiver type	GALILEO	E1B/C E5b
	BEIDOU	B1I B2I
	QZSS	L1C/A L1S L2C
	SBAS	L1C/A
Position accuracy	RTK	0.01 m + 1 ppm CEP
Convergence time	RTK	< 10 sec
	Cold starts	24 s
Acquisition	Aided starts	2 s
	Reacquisition	2 s
	Tracking & Nav.	-167 dBm
	Cold starts	-148 dBm
Sensitivity	Hot starts	-157 dBm
	Reacquisition	-160 dBm
	NMEA	
	UBX binary	
Protocols	RTCM v. 3.3	
	SPARTN v. 2.0	
	54-pin LGA (land grid array)	
Package	17 x 22 x 2.4 mm	

Figure 5. ZED-F9P module features [17]

Figure 6. Low-cost and geodetic receivers placed on an apparatus

13 days of observation data were collected for both receivers. During the data collection, the cut-off angle was chosen as 10 degrees. GPS+GLONASS+Galileo+BeiDou-2+BeiDou-3 (GREC2C3) combination was used for u-blox receiver, GPS+GLONASS (GR) combination was used for the geodetic receiver. Observation data were collected on days when the number of satellites was sufficient, in clear sky conditions and with good geometry. The number of GPS, GLONASS, Galileo and BeiDou satellites on the fifth day is given in Fig. 7. Also, there is no effect that will cause multipath effect in the area where both receivers are located. The location information of the station is 37° 51' 58.5109" N and 32° 25' 9.1874" E.
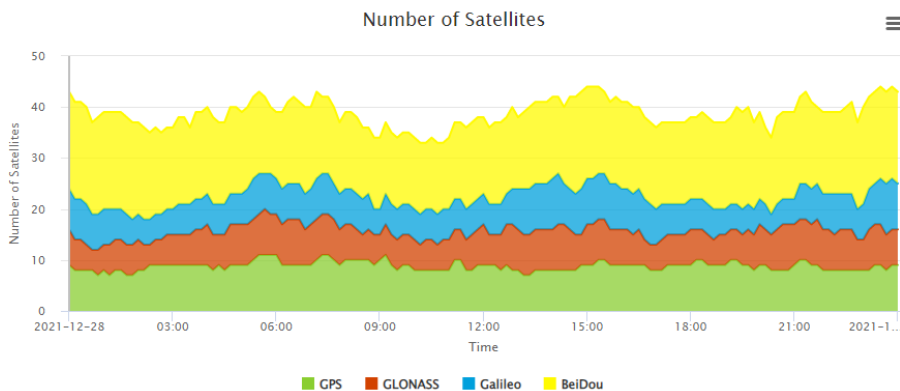
Figure 7. GPS, GLONASS, Galileo and Beidou number of satellites [18]

The test data has been examined, and the required software has been built for the low-cost GNSS module and antenna kit to be ready for measurement (such as converting raw data to RINEX and setting measurement settings).

For remote access to raspberry pi devices to which low-cost GNSS receivers are connected, internet connection infrastructure has been installed on the roof. In this way, the RINEX data recorded by the low-cost GNSS receiver can be accessed via the remote access program (WinSCP software). Raw data of the low-cost GNSS receiver was converted to RINEX data using convbin.exe software. The raspberry pi and screen-keyboard set devices connected to U-blox receiver board are given in Fig. 8 and Fig. 9.

Figure 8. Raspberry pi 4 model B

Figure 9. Raspberry pi 4 monitor and keyboard kit

FINDINGS

Observation data obtained for 13 days (24 December 2021-05 January 2022) with low-cost and geodetic receivers were converted into RINEX format. Then, using the Linux shell script, the epoch availability of the RINEX observation files for each day was determined. Epoch availability percentages of geodetic and low-cost receivers are provided in Table 1. In addition, its graphic representation is given in Fig. 10. As depicted in Table 1 and Fig. 10, the epoch availability percentage of the geodetic receiver is mostly 100%. Only the results obtained on 27 December 2021 and 28 December 2021 are below this value, and the epoch availability percentages are 99.86% and 99.93%, respectively. When the results of the low-cost GNSS receiver are examined, it is seen that the epoch availability for 9 days is 100%. The minimum epoch availability percentage was obtained as 94.67% on 04 January 2022. In addition, epoch availability percentages of 99.65%, 98.54%, and 97.07% were obtained on 26 December 2021, 01 January 2022, and 05 January 2022, respectively.

Table 1. Epoch availability percentage of geodetic and low-cost receivers

Days	Geodetic Receiver (%)	Lox-Cost Receiver (%)
24.12.2021	100	100
25.12.2021	100	100
26.12.2021	100	99.65
27.12.2021	99.86	100
28.12.2021	99.93	100
29.12.2021	100	100
30.12.2021	100	100
31.12.2021	100	100
1.01.2022	100	98.54
2.01.2022	100	100
3.01.2022	100	100
4.01.2022	100	94.67
5.01.2022	100	97.07
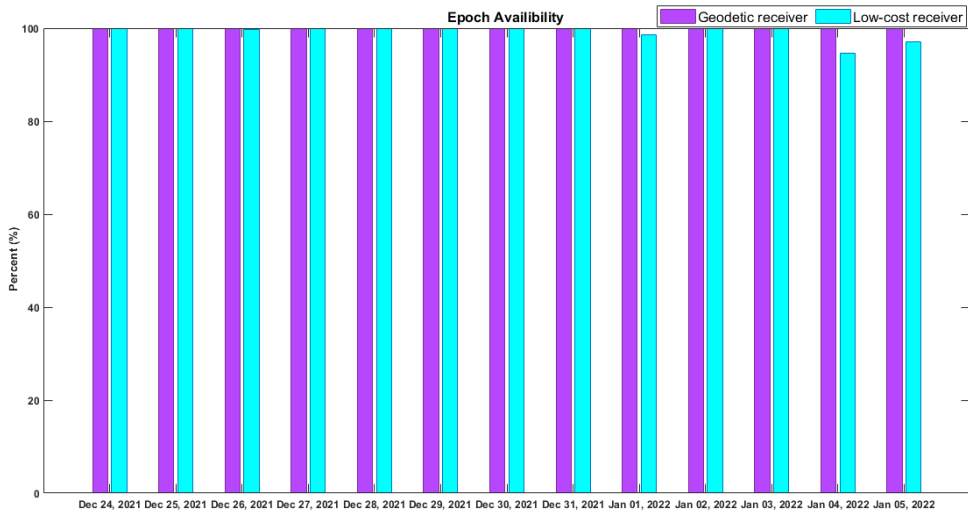
Figure 10. Epoch availability percentage of geodetic and low-cost receivers

RESULTS

In this study, the epoch availability performance of the low-cost GNSS receiver was tested with a geodetic GNSS receiver in a comparative approach. For this purpose, 13 days of RINEX observation data were used. The results obtained showed that the epoch availability percentages of the observation files based on geodetic GNSS were mostly 100%, and over 99% on the other days. When the low-cost GNSS receiver was used, 100% epoch availability percentages were generally achieved, and minimum epoch availability was obtained as 94.67%. The overall results showed that the low-cost GNSS receiver is good enough in terms of epoch availability.

REFERENCES

[1] N. Manzini *et al.*, 'Performance analysis of low-cost GNSS stations for structural health monitoring of civil engineering structures', *Struct. Infrastruct. Eng.*, vol. 18, no. 5, pp. 595–611, 2022.

[2] L. Lu, L. Ma, T. Wu, and X. Chen, 'Performance analysis of positioning solution using low-cost single-frequency u-blox receiver based on baseline length constraint', *Sensors (Switzerland)*, vol. 19, no. 19, Oct. 2019.

[3] L. Poluzzi, L. Tavasci, F. Corsini, M. Barbarella, and S. Gandolfi, 'Low-cost GNSS sensors for monitoring applications', *Appl. Geomatics*, vol. 12, pp. 35–44, Apr. 2020.

[4] P. J. Knight, C. O. Bird, A. Sinclair, and A. J. Plater, 'A low-cost GNSS buoy platform for measuring coastal sea levels', *Ocean Eng.*, vol. 203, May 2020.

[5] M. S. Garrido-Carretero, M. C. de Lacy-Pérez de los Cobos, M. J. Borque-Arancón, A. M. Ruiz-Armenteros, R. Moreno-Guerrero, and A. J. Gil-Cruz, 'Low-cost GNSS receiver in RTK positioning under the standard ISO-17123-8: A feasible option in geomatics', *Meas. J. Int. Meas. Confed.*, vol. 137, pp. 168–178, Apr. 2019.

[6] V. Hamza, B. Stopar, T. Ambrožič, and O. Sterle, 'Performance evaluation of low-cost multi-frequency gnss receivers and antennas for displacement detection', *Appl. Sci.*, vol. 11, no. 14, Jul. 2021.

[7] L. Tunini, D. Zuliani, and A. Magrin, 'Applicability of Cost-Effective GNSS Sensors for Crustal Deformation Studies', *Sensors*, vol. 22, no. 1, Jan. 2022.

[8] N. Wielgocka, T. Hadas, A. Kaczmarek, and G. Marut, 'Feasibility of using low-cost dual-frequency gnss receivers for land surveying', *Sensors*, vol. 21, no. 6, pp. 1–14, Mar. 2021.

[9] R. Romero-Andrade, M. E. Trejo-Soto, A. Vega-Ayala, D. Hernández-Andrade, J. R. Vázquez-Ontiveros, and G. Sharma, 'Positioning evaluation of single and dual-frequency low-cost gnss receivers signals using ppp and static relative methods in urban areas', *Appl. Sci.*, vol. 11, no. 22, Nov. 2021.

[10] Y. Zhang *et al.*, 'Static and kinematic positioning performance of a low-cost real-time kinematic navigation system module', *Adv. Sp. Res.*, vol. 63, no. 9, pp. 3029–3042, May 2019.

[11] D. J. Purnell, N. Gomez, W. Minarik, D. Porter, and G. Langston, 'Precise water level measurements using low-cost GNSS antenna arrays', *Earth Surf. Dyn.*, vol. 9, no. 3, pp. 673–685, Jun. 2021.

[12] Z. Nie, F. Liu, and Y. Gao, 'Real-time precise point positioning with a low-cost dual-frequency GNSS device', *GPS Solut.*, vol. 24, no. 1, Jan. 2020.

[13] R. Odolinski and P. J. G. Teunissen, 'An assessment of smartphone and low-cost multi-GNSS single-frequency RTK positioning for low, medium and high ionospheric disturbance periods', *J. Geod.*, vol. 93, no. 5, pp. 701–722, May 2019.

[14] R. Odolinski and P. J. G. Teunissen, 'Best integer equivariant estimation: performance analysis using real data collected by low-cost, single- and dual-frequency, multi-GNSS receivers for short- to long-baseline RTK positioning', *J. Geod.*, vol. 94, no. 9, Sep. 2020.

[15] D. Notti *et al.*, 'Low-cost GNSS solution for continuous monitoring of slope instabilities applied to Madonna del Sasso Sanctuary (NW Italy)', *Sensors (Switzerland)*, vol. 20, no. 1, 2020.

[16] "Survey GNSS Multiband antenna (IP66)." Ardusimple https://www.ardusimple.com/product/survey-gnss-multiband-antenna (accessed August 20, 2022)

[17] "ZED-F9P-Module." Ublox. https://www.u-blox.com/en/product/zed-f9p-module (accessed August 20, 2022)

[18] "Trimble GNSS Planning Online" GNSS Planning https://www.gnssplanning.com/#/charts (accessed August 25, 2022)

NONLINEAR SLOSHING RESPONSE OF LIQUID-FILLED REINFORCED CONCRETE (RC) ELEVATED WATER TANK UNDER SEISMIC EXCITATION

Olgun KÖKSAL[1], Zeki KARACA[2], Erdem TÜRKELİ[3]

INTRODUCTION

Drinking water tanks, which serve for daily use, meet the necessary water needs without being damaged in disasters such as earthquakes and floods. It is known that these warehouses have different load carrying systems depending on their location on the ground and the type of material. Reinforced concrete (RC) elevated water tanks are built to store drinking water in areas where the water pressure in the network is insufficient. RC elevated water tanks are more affected by earthquakes than other types of water tanks. For this reason, earthquake risk analysis of the region where the RC elevated water tank is located should also be carried out appropriately. Scientific studies on liquid tanks started at the end of the 19th century and are still continuing. By examining the scientific literature, the following studies have been carried out on elevated water tanks: Studies in which the water mass in the tank is considered as a single mass and the oscillation is neglected [1], studies using the multi-mass system approach [2], [3], studies considering the added mass approach [4]. The pressure acting on the tank wall in static state is expressed as hydrostatic pressure. However, earthquake-induced ground movements shake the water in the tank. In this case, the pressure exerted by the water on the wall is called as the hydrodynamic pressure. The water in the tank oscillates and behaves differently from the main load carrying system. While making dynamic calculations of elevated water tanks, hydrodynamic pressures should be determined as close to real results as possible. For this, there are many methods such as Westergaard's method [5], Hoskins and Jacobsen's method [6], Werner-Sundguist method [7], Housner's method [2], Haroun's method [1] which makes the distribution of hydrodynamic pressure on the wall surface according to various assumptions. There are also Finite Element Method and Finite Volume Methods.

MATERIAL AND METHOD

As mentioned earlier, there are many methods for finding hydrodynamic pressures. One of these methods is the Smoothed Particle Hydrodynamics (SPH)

1   Samsun University, Vocational School, Construction Department, Samsun/TURKEY, olgun.koksal@samsun.edu.tr, (ORCID: https://orcid.org/0000-0001-8448-6404)
2   Ondokuz Mayis University, Faculty of Engineering, Department of Civil Engineering, Samsun/TURKEY, zekikrc@omu.edu.tr, (ORCID: https://orcid.org/0000-0003-1130-5259)
3   Ordu University, Vocational School of Technical Sciences, Construction Department, Ordu/ TURKEY,erdemturkeli@odu.edu.tr, (ORCID: https://orcid.org/0000-0002-4293-4712)

method [8]. SPH is a meshless method based on the Lagrangian approach, which is widely used in computational fluid mechanics, where the flow of particles is represented and can interact with structures, and large deformations are obtained. The first use of this method is encountered by Gingold, Monaghan [9], and Lucy [10] in solving astrophysics problems. In addition, many researchers have used the SPH method [11-16]. It is aimed to solve the numerical equations of fluid dynamics by accepting the fluid as a particle with SPH. Since the fluid in the desired geometry can be modeled with SPH, it is preferred for fluid modeling in many areas today. Various particle clusters are organized by determining an interpolation function (kernel function) [17] based on the SPH Langrange approach.

Using the Navier-Stokes equations in computational fluid dynamics, the position of each particle in the fluid model can be integrated according to the physical properties around the radius r. A neighboring or contiguous array of particles surrounding r is determined by the interpolation length "h" combined at either a two-dimensional circular or a three-dimensional spherical function-based distance. In each time step of the determined period, new physical values are calculated and transferred to the updated new values (Fig. 1).

Using the kernel function, the Navier-Stokes momentum equations, continuity equation and equation of state are calculated by weight of any particle in the determined particle set for each time step. Since the conservation equations of the Kernel function are expected to give possible values depending on the interpolation function at any point, this function is converted into a differential form in accordance with the particle. Thus, the function can simulate the particle's displacement over time.
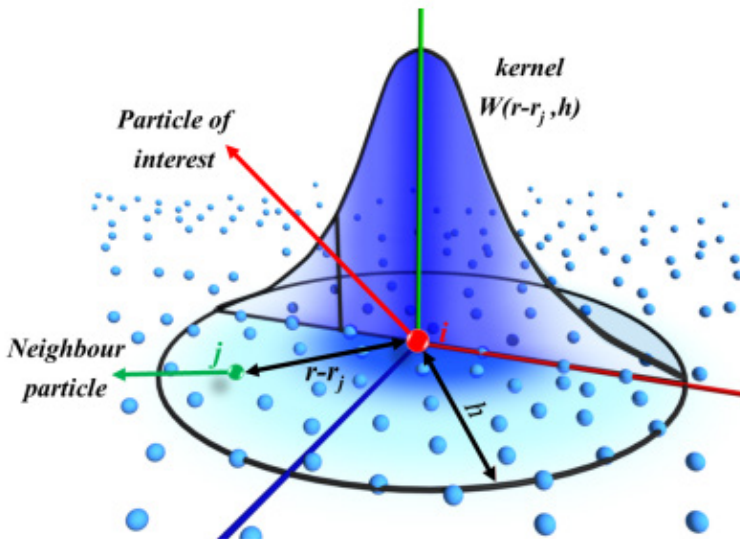
Figure 1: W(r) The principle of the SPH kernel function [18].

FORMULATION

Partial differential equations can be used if it is desired to define the flow problem in a nonlinear way in fluid dynamics. Since the Lagrangian approximation is based on the SPH, which is independent of the meshing, a certain volume of fluid can be modeled with the assumption of a finite number of particles. These particles can carry both their own mass and any physical properties. The following integral interpolation is used to express the motion of the particle [11,14].

$$F(r) = \int_\Omega F(r')W(r - r', h)d \tag{1}$$

In Eq.(1), r is denoting the location where the variable F is obtained, r' represents the location where the variable is known, Ω is the solution area, W is the weight function, h is the action distance of the weight function. The expression $W(r - r')$ is known as the Kernel function mentioned above. The function F can be approximated in a discontinuous, discrete way based on the set of particles. In this case, the function is interpolated on a particle "a" and becomes a sum over all particles of interpolation length h [11, 14].

$$F(r_a) \approx \sum_b F(r_b)W(r_a - r_b, h)\Delta v_b \tag{2}$$

v_b indicates the volume of the neighboring particle and $v_b = \dfrac{m_b}{\rho_b}$ where m_b represents the mass of the particle and ρ_b is its density. In this manner, Eq.(2) becomes;

$$F(r_a) \approx \sum_b F(r_b)\frac{m_b}{\rho_b}W(r_a - r_b, h) \tag{3}$$

The accuracy of a model prepared by the SPH method depends on the good selection of the Kernel function [17]. In this study, Eq.(4) is used as a kernel function

$$(\ ) = \alpha \left(1 - \frac{\ }{2}\right)^4 (2\ + 1) \qquad 0 \leq\ \leq 2 \tag{4}$$

In three dimensions, $\alpha_D = \dfrac{21}{164\pi h^3}$. Also, q is the dimensionless distance between particles a and b according to the given r/h ratio. Conservation of momentum in a continuous medium can be expressed as Eq.(5) [18].

$$\frac{dv}{dt} = -\frac{1}{\rho}\nabla P + g + \Gamma \tag{5}$$

where, Γ and g are denoting viscous force and gravitational acceleration, respectively.

The momentum equation is used to describe the acceleration of a particle as a result of interaction with a neighboring particle.

Momentum equation in SPH method can be given by Eq.(6) [19]

$$\frac{dv_a}{dt} = -\sum_b m_b \left(\frac{P_a + P_b}{\rho_a + \rho_b} + \Pi_b \right) \nabla_a W_b + g \tag{6}$$

where Π_{ab} is denoting the viscous terms between particles a and b which is given in Eq.(7)

$$\Pi_{ab} = \begin{cases} \dfrac{-\alpha \overline{c_b}\, \mu_b}{\rho_b} & v_b \cdot r_b < 0 \\ 0 & v_b \cdot r_b > 0 \end{cases} \tag{7}$$

where $r_{ab} = r_a - r_b$ is denoting particle coordinates, $v_{ab} = v_a - v_b$ is the viscosity, $\overline{c_{ab}} = 0.5(c_a + c_b)$ is average sound velocity of the particles, $\eta^2 = 0.01\, h^2$ and α (a 0.01) are suitable coefficients to provide the propagation.

Continuity equation for a weakly compressible fluid [19] can be expressed by Eq.(8).

$$\frac{d\rho_a}{td} = \rho_a \sum_b \frac{m_b}{\rho_b} v_{ab} \nabla_a W_{ab} \tag{8}$$

Conservation of mass is expressed in this equation by the SPH method. Since the fluid is considered weakly compressible in the SPH method, the pressure value of the fluid is calculated with the help of the equation of state [20].

$$P = b \left[\left(\frac{\rho_a}{\rho_0} \right)^{\gamma} - 1 \right] \tag{9}$$

b is a constant value and is related to the modulus of elasticity of the fluid. $b = c_0^2 \dfrac{\rho_0}{\lambda}$ $\rho_0 = 1000 kg/m^3$ is the specific mass of water. γ can take values from 1 to 7 ($\gamma = 7$). The expression -1 in the equation ensures that the pressure on the fluid free surface is zero.

NUMERICAL APPLICATION

The water pressures acting to the walls as a result of an earthquake ground motion acting on the RC elevated water tank can be calculated by the SPH method. Using the DualSphysics program [19] based on this method, hydrodynamic pressures on the walls can be obtained. Fig.2 shows Flow-chart of DualSPHysics program.
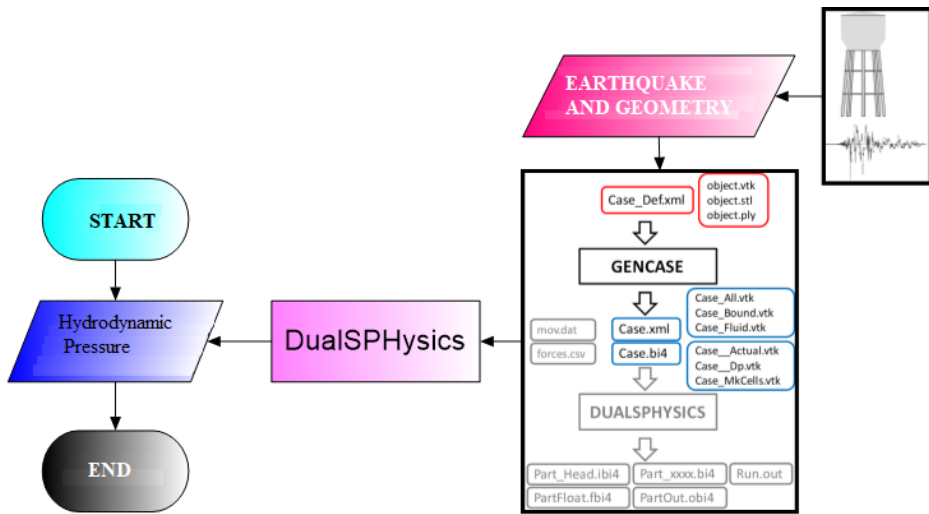
Figure 2: Flow-chart of DualSPHysics program.

This program performs analysis with the computer's Central Processing Unit (CPU) and Graphics Processing Unit (GPU). Due to the large number of particles and visual presentations, it is more convenient to work on the GPU in terms of time. For this purpose, a solid model of the 1000 m³ RC elevated water tank was produced. Then, with the help of Matlab Partial Differential Equation toolbox [21], the load carrying system of the cited elevated water tank was arranged and the finite element mesh of the RC part was formed. Considering the half-full and full-filled state of the tank, structural analyzes were carried out in the DualSphysics program [19].
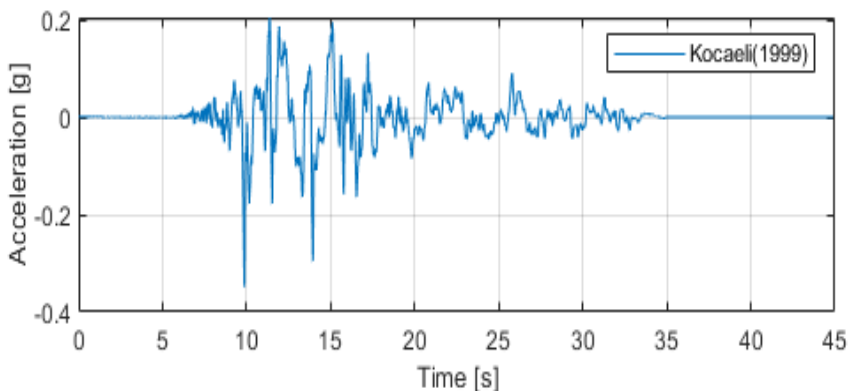
Figure 3: Acceleration time histories of Kocaeli earthquake (YPT330), 1999 [22].

The Kocaeli earthquake, which took place in Turkey on August 17, 1999, was chosen for the earthquake ground motion. The greatest ground acceleration of this earthquake was 0.349g in 9.87 seconds. Fig. 3 shows the acceleration-time variation of the cited Kocaeli earthquake.
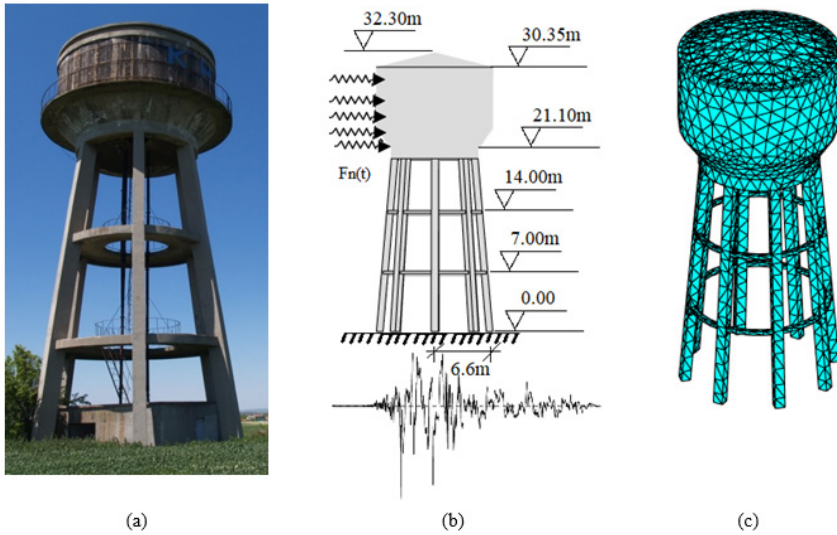
Figure 4: (a) Photograph, (b) mathematical model and (c) finite element model of chosen RC elevated water tank.

Table 1: Material properties of reinforced concrete elevated water tank.

Material	Modulus of Elasticity (GPa)	Poisson Oranı	Mass Density (kg/m³)
Concrete	30	0.2	2400

Fig. 4 shows the photo (a), mathematical model (b), and finite element model (c) of the RC elevated water tank. Table 1 shows material properties of reinforced concrete elevated water tank.

As can be seen from the figure, the load carrying system of the tank consists of a cylinder and a truncated cone, eight columns and two beams. Columns have 83.5cmx75cm rectangular section, beams have 35cmx35cm square section, and the thickness of the tank wall is 30cm. The water volume of the tank is 1000m³. The water level of the tank from ground at 100% full level is 30.35m, the water height is 9.25m, the water level of the tank from ground at 50% full level is 25.73m, and the water height is 4.63m. Solid tetrahedral elements are used for the finite element. In the finite element model, there are 3894 nodes and 11051 finite elements. The number of finite element nodes on the tank walls is 2199. Water volume at 50% full and 100% full level was analyzed according to Kocaeli earthquake acceleration records with the DualSPHysics program based on the SPH method. Fig. 5 shows solid tetrahedral finite element. Table 2 shows finite element properties of reinforced concrete elevated water tank.
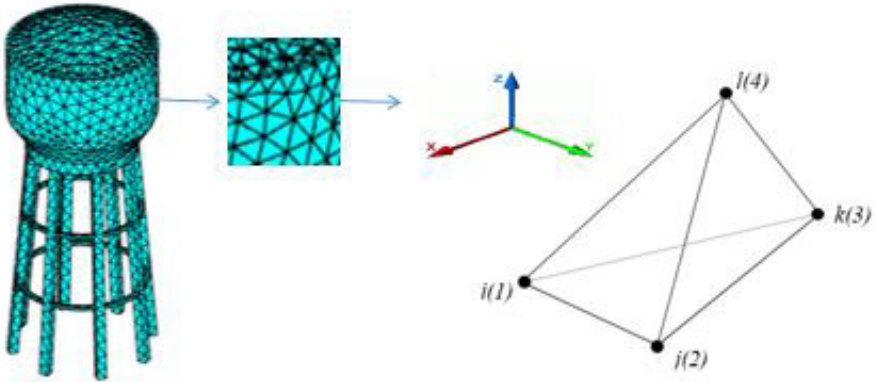
Figure 5: Solid tetrahedral finite element.

Table 2: Finite element properties of reinforced concrete elevated water tank.

Nodes	Elements	Maximume Element Size	Minimume Element Size	Mesh Gradation	Geometric Order
3x3894=11682	4x11051=44204	1.4993	0.7496	1.5	Linear

3.1. Case 1: The tank is half full (50 %)

Table 3 shows the analysis parameters required for the cited RC elevated water tank to be 50% full.

Table 3: Parameters of the SPH method for half-filled tank.

Property	Options
Version	DualSPHysics v4.056
Dimension	3D
Type of the kernel function	Wendland
Time-stepping	Verlet
Density filter	Shepard filter
Viscosity treatment	Artificial viscosity (α=0.01)
Boundary conditions(BC)	Dynamic
Distance between particle(dp)	0.09m
Smoothing length	1
The number of fluid particles	38029
Density of water	1000kg/m³
CFL coefficient	0.20
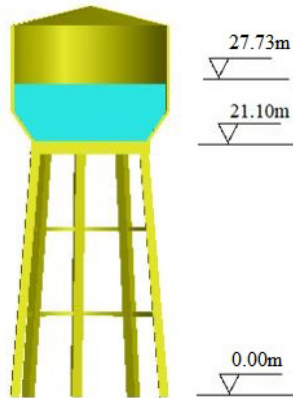
Figure 6: Case I, the elevated water tank is half full.

Also, Fig. 6 shows the state of the water at 50% full. At this level, the water free surface elevation is 25.73 m (from ground), the water height is 4.63 m and the volume of water is 528.41 m^3.

3.2. Case 2: The tank is full (100 %)

Table 4 shows the analysis parameters required for the cited RC elevated water tank to be 100% full. Also, Fig. 7 shows the state of the water at 100% full.
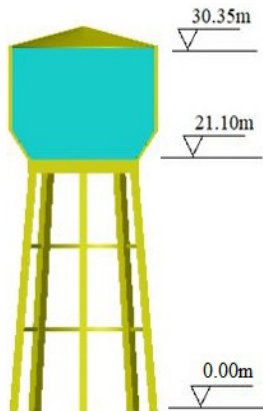
Figure 7: Case II, the elevated water tank is full.

Table 4: Parameters of the SPH method for full-filled tank.

Property	Options
Version	DualSPHysics v4.056
Dimension	3D
Type of the kernel function	Wendland
Time-stepping	Verlet
Density filter	Shepard filter
Viscosity treatment	Artificial viscosity (α=0.01)

Boundary conditions(BC)	Dynamic
Distance between particle(dp)	0.20m
Smoothing length	1
The number of fluid particles	95333
Density of water	1000kg/m³
CFL coefficient	0.20

At this level, the water free surface elevation is 30.35 m (from ground), the water height is 9.25 m and the water volume is 1000 m³.

RESULTS

As a result of the analyses carried out, all hydrodynamic pressures at the finite element nodal points on the tank wall were obtained nonlinearly. As an example, the time dependent hydrodynamic pressure change of a point at the bottom of the tank at X=5.814 m, Y=11.396 m, Z=21.1 m coordinates is given in Figs 8 and 9 for half (50%) and complete full state (100%), respectively.

4.1. Case 1: The tank is half full (50 %)
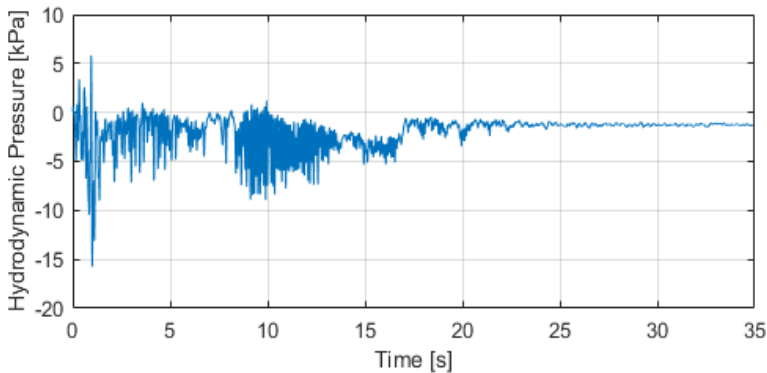
Figure 8: Case I, the elevated water tank is half full (50%).
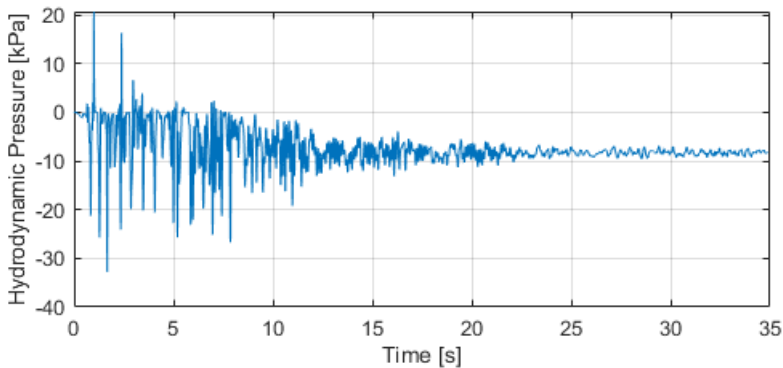
Figure 9: Case II, the elevated water tank is full (100%).

From the interpretation of Figs 8 and 9, these following results can be obtained. In Fig. 8, the hydrodynamic pressure is 0.59 kPa at 9.87sec when the greatest ground acceleration occurs. In Fig. 9, the hydrodynamic pressure is 9.29 kPa at 9.87sec when the greatest ground acceleration occurs.

CONCLUSION

In this study, the behavior of the water in the RC elevated water tank under the effect of earthquake ground motion was investigated for half and full conditions. The SPH method [11,14] is a meshless method used in nonlinear fluid modeling in computational fluid dynamics. Using this method, the nonlinear hydrodynamic pressures on the wall of the water tank under earthquake effect were calculated. Analyzes were carried out for the case of the elevated water tank being half (50%) and full (100%) by using the DualSPHysics program[19]. As a result of the analyses, the hydrodynamic pressures at any point on the tank wall were obtained. As an example, the time dependent hydrodynamic pressure change of a point on the wall at the coordinates X=5.8814 m,Y=11.396 m, Z=21.1 m was investigated. As a result of this examination, it was observed that the value of the hydrodynamic pressure when the tank is half full is approximately 6.35% of the hydrodynamic pressure when the tank is full (half full [50%] is approximately 20 times smaller than complete full [100%]).

REFERENCES

[1] M. A. Haroun and H. Ellaithy, "Seismically Induced Fluid Forces On Elevated Tanks.," J. Tech. Top. Civ. Eng., vol. 111, no. 1, pp. 1–15, Dec. 1985, doi: 10.1061/JTCEDL.0000023.

[2] G. H. the seismological society of America and U. 1957, "Dynamic pressures on accelerated fluid containers," pubs.geoscienceworld.org.

[3] R. Livaoglu and A. Dogangun, "Effect of foundation embedment on seismic behavior of elevated tanks considering fluid-structure-soil interaction," Soil Dyn. Earthq. Eng., vol. 27, no. 9, pp. 855–863, Sep. 2007, doi: 10.1016/j.soildyn.2007.01.008.

[4] A. Dogangun, A. Durmus, and Y. Ayvaz, "Earthquake analysis of flexible rectangular tanks by using the Lagrangian fluid finite element," Eur. J. Mech. A, Solids, vol. 16, no. no.1, pp. 165–182, 1997.

[5] H. M. Westergaard, "Water Pressures on Dams during Earthquakes," Trans. Am. Soc. Civ. Eng., vol. 98, no. 2, pp. 418–433, Jan. 1933, doi: 10.1061/TACEAT.0004496.

[6] L. Hoskins, L. J.-B. of the Seismological, and U. 1934, "Water pressure in a tank caused by a simulated earthquake," pubs.geoscienceworld.org.

[7] P. W. Werner and K. J. Sundquist, "On hydrodynamic earthquake

effects," Eos, Trans. Am. Geophys. Union, vol. 30, no. 5, pp. 636–657, 1949, doi: 10.1029/TR030I005P00636.

[8] J. M.-J. of computational physics and U. 1994, "Simulating free surface flows with SPH," Elsevier.

[9] R. Gingold, J. M.-M. notices of the Royal, and U. 1977, "Smoothed particle hydrodynamics: theory and application to non-spherical stars," academic.oup.com.

[10] L. L.-T. astronomical journal and U. 1977, "A numerical approach to the testing of the fission hypothesis," adsabs.harvard.edu.

[11] J. M.-A. review of astronomy and astrophysics and U. 1992, "Smoothed particle hydrodynamics," adsabs.harvard.edu.

[12] A. Crespo, … M. G.-G.-C., and undefined 2007, "Boundary conditions generated by dynamic particles in SPH methods," scholars.northwestern.edu.

[13] M. Gomez-Gesteira, B. D. Rogers, R. A. Dalrymple, and A. J. C. Crespo, "State-of-the-art of classical SPH for free-surface flows," J. Hydraul. Res., vol. 48, no. SUPPL. 1, pp. 6–27, 2010, doi: 10.1080/00221686.2010.9641242.

[14] G. Liu and M. Liu, "Smoothed particle hydrodynamics: a meshfree particle method," 2003.

[15] M. Kusić, J. Radnić, N. Grgić, A. H.- Građevinar, and U. 2018, "Sloshing in medium size tanks caused by earthquake studied by SPH," pdfs.semanticscholar.org, vol. 8, pp. 671–684, 2018, doi: 10.14256/JCE.2169.2017.

[16] A. E. Güray, A. G. Yazıcı, E. Güray, and G. Yazıcı, "Sloshing in medium sized tanks under earthquake load," hrcak.srce.hr, vol. 7, pp. 655–662, 2015, doi: 10.14256/JCE.1149.2014.

[17] H. Wendland, "Piecewise polynomial, positive definite and compactly supported radial functions of minimal degree," Adv. Comput. Math., vol. 4, no. 1, pp. 389–396, Dec. 1995, doi: 10.1007/BF02123482.

[18] R. Sampath, N. Montanari, N. Akinci, S. Prescott, and C. Smith, "Large-scale solitary wave simulation with implicit incompressible SPH," J. Ocean Eng. Mar. Energy, vol. 2, no. 3, pp. 313–329, Aug. 2016, doi: 10.1007/S40722-016-0060-8.

[19] "SPHYSICS Home Page -SPHYSICS." https://wiki.manchester.ac.uk/sphysics/index.php/Main_Page (accessed Oct. 28, 2022).

[20] R. Ibrahim, "Liquid sloshing dynamics: theory and applications," 2005.

[21] "Partial Differential Equation Toolbox - MATLAB." https://www.mathworks.com/products/pde.html (accessed Oct. 28, 2022).

[22] "Pacific Earthquake Engineering Research Center." https://peer.berkeley.edu/smcat/data (accessed Oct. 30, 2022).

STRUCTURAL AND FIRE PERFORMANCE OF COMPOSITE SLAB SYSTEMS PROTECTED BY INTUMESCENT COATING

Burak Kaan CIRPICI[1], Melih ERMANCIK[2]

INTRODUCTION

Steel sheets have been used as formwork in conventional reinforced concrete floors since the 1920s, and starting in the 1950s, concrete was commonly used as tension reinforcement in composite floor structural elements. The aforementioned steel sheets' resulting grooves and protrusions were the first prototypes of contemporary composite flooring in the 1960s. Due to improvements in the materials used to create the composite flooring system throughout time as well as changes in design standards, the building components are lighter and more economical.

The first instances of composite floors appeared as a result of the initial usage of steel sheet as a permanent concrete formwork for floors of reinforced concrete structures. At various points in time, these elements were used as various elements. These components served a variety of purposes at various times. With the introduction of many sheet types, every nation started to create its own standards. For instance, in 1967, the Iron and Steel Institute of America began studying mixed records. These studies served as the foundation for American, British, and European standards for composite design systems [1-4]. Shear bond rupture frequently occurs when fatigue loads greater than the bond load between the steel beam and shear studs are applied in composite floors [5, 6]. Bailey and Moore [7] studied a novel approach to the construction of steel-framed buildings with composite floor slabs exposed to fire. Fire tests were conducted as part of the study in a real steel-framed building. According to tests, the performance of composite flooring systems is undervalued as compared to the conventional design processes for the fire boundary issue. In their study, Üstündağ and Çelik [8] provided information regarding composite flooring systems used in multi-story steel structures. In the study, new methods are described and recommendations are offered in light of technological advancements from 2006. In his MSc thesis, Yılmaz [9] looked at the positive moment carrying capacity of composite flooring parts. Six simple beam composite slab samples with a similar sheet thickness and a slab height of 4.10 m were evaluated in the study to see how the impact of vertical loads changed the trapezoidal sheet metal elements utilized

1   Affiliation : Erzurum Technical University, Engineering and Architecture Faculty, Civil Engineering Department, Erzurum, TURKEY
2   Affiliation : Erzurum Technical University, Engineering and Architecture Faculty, Civil Engineering Department, Erzurum, TURKEY

in the slab. Data from experiments and theory were contrasted. Three approaches to designing a composite steel-concrete slab were taken into consideration and used in a metro station in the study by Cordeiro and Silva [10]. It took 90 minutes for the mezzanine floor of the subway to catch fire. Steel beams served as a fire barrier. The membrane effect was modeled using the streamlined Bailey approach and VULCAN software based on the findings. Selamet and Yolaçan [11] used a symmetrical joint mechanism on both sides of the composite floor created for a high-rise steel building and gave a bigger floor area test opportunity than the experimental area in their experimental investigation (furnace). They proposed that the secondary (secondary) steel beams in composite floor slabs remain uninsulated after it was found that the concrete slab put through the test demonstrated membrane behavior during fire and loading and maintained its load-bearing property throughout the fire. By figuring out the critical temperature and fire protection values of the composite beams, which are already included in the Chinese Technical Standard for fire safety of steel structures, Li and Wang [12] developed a simplified approach based on the critical temperature for the fire resistance of steel-concrete composite beams. A strategy has been suggested. When the steel sheet is not directly exposed to the fire, the temperature on the exposed top surface is especially crucial [13-17]. If, for some reason, the fire spreads to the top surface of the concrete, this surface may start to take on greater significance, leading to the formation of fractures in the concrete block and the spread of the fire to steel components, such as the slab and steel. This topic was covered by Wang, et al. [18], and scholars are still interested in it. In order to study slab behavior in terms of deflection and load bearing capacity with fire testing and numerical models, Nguyen, et al. [13] built a small-scale partially shielded composite floor with unprotected secondary and interior beams and shielded edge beams. They came to the conclusion that the inner beams may be left exposed because rotational restriction along the covered edge beams created severe stress concentration on these beams in the slab and no structural damage could be observed. Based on fire tests and loaded numerical simulations, Nguyen and Tan [19] evaluated the impact of the bending stiffness of protected edge beams, including the relationship between element temperatures and mid-span deflections. The rigidity of the protected primary beams can be increased in order to limit slab deflection, and the composite movement between the beams is crucial to this advantageous outcome. The behavior of transverse trapezoidal sheet metal and head shear nails fastened to the flooring was examined in both ambient and fire situations in order to covertly discuss the Eurocode design approach, ascertain the capacity of the shear joint, and offer a new design formula. By using computational three-dimensional models to describe the non-linear behavior of steel beams independent of the kind of slab with shear studs, the rotational capacity of steel I-section beams at increasing temperatures was explored Pantousa and Mistakidis [20]. Alam, et al. [21] sought to analyze the fire behavior of unprotected and protected thin slabs using a finite

element analysis approach. Their findings show that protected thin floors provide greater fire resistance since the steel temperature stays below 400 °C after a regular 60-minute fire exposure. In their study, Jiang, et al. [22] found that while welded wire mesh reinforcement is not always required in composite flooring elements, it does aid in crack prevention. Lim and Wade [23] also noted the buckling of the steel sheet subjected to direct fire, the separation of the steel sheet from the concrete with the effect of loading at high temperatures, and the pouring of the concrete due to the heat penetration due to the cracks generated. After 20 minutes of fire testing in this investigation, cracks began to spread from the point where the top of the slit and the concrete slab met at an angle of roughly 45 degrees. Cirpici, et al. [17] carried out heat transfer analyses on the slab by igniting Fast and Standard (ISO 834) fires from the top of the concrete component for a period of one hour without loading the composite slab. In the computational model they made, they added fire protection to the steel profile and sheet metal and figured out how the temperature was distributed in the floor's structural parts. In recent years, Piloto, et al. [24], Piloto, et al. [25], and Piloto, et al. [26] have conducted numerical studies on the bond separation effect between concrete and steel blocks due to rapid heating of steel profiles during exposure to a fire in order to develop a new simple calculation model to describe the fire resistance of composite boards, taking into account the load level and bond separation.

In this study, the span length (long or short) and whether the sections are fire-protected or not are identified as potential parameters that affect the performance of composite flooring under fire. It has been investigated to what extent these characteristics affect the performance of a composite floor under Standard (ISO 834) fire conditions. Using Eurocode design equations and the Excel application (Macros) to conduct analyses, a new mathematical model was constructed. A passive fire protection material (intumescent coating) was placed on the steel beam beneath the composite floor, and beam temperature analyses were conducted at different beam sections. As a result of these analyses, it was determined how the mechanical properties of the steel change at elevated temperatures. In addition, the displacements under various loading circumstances have been determined.

MATERIAL AND METHOD

Temperature Predictions in Steel Beams

Unprotected Steel

$$\Delta T_{st} = k_{sh} \frac{A_m/V}{\rho_{st} C_{st}} \frac{A_s}{V} \dot{h}_{net} \Delta t \tag{1}$$

k_{sh} - Correction factor for shading effect

A_m/V – Section factor (1/m)

A_m – Surface area of a unit length element (m2/m)

V – Volume of structural element unit length (m3/m)

C_{st} – Specific heat of steel (J/kgK)

ρ_{st} – unit volume mass of steel

$\dot{h}_{net}$ – Design value of net heat axis per unit area (W/m2)

Δt – Time step (sec)

Correction factor for shadow effect in I – sections under the influence of standard fire (ISO 834); $k_{sh} = 0.9[A_m/V]_b/[A_m/V]$

In all other cases; $k_{sh} = [A_m/V]_b/[A_m/V]$

$\dot{h}_{net} = \varepsilon_f \varepsilon_m$ where $\varepsilon_f = 1.0$ and ε_m should be taken from EN 1991-1-2 [27]

Protected Steel

According to EN 1993-1-2, the temperature of a protected steel section is calculated using Equation below.

$$\Delta T_{st} = \frac{\lambda_p A_p/V}{d_p c_{st} \rho_{st} \left(1 + \frac{\phi}{3}\right)} \frac{T_f - T_{st}}{} \Delta t - \left(e^{\phi/10} - 1\right) \Delta T_f \tag{2}$$

with $\phi = \frac{c_p \rho_p}{c_{st} \rho_{st}} d_p \frac{A_p}{V}$

where λ_p (W/mK) is the thermal conductivity of the fire protection material, A_p/V (m^{-1}) is the section factor of the protected steel section based on the diameter, d_p (m) is the fire protection thickness, c_p (J/kgK) and ρ_p (kg/m^3) are specific heat and density of the protection material, T_f $(°C)$ and T_{st} $(°C)$ is the exposed fire temperature (ISO) and steel temperature respectively, Δt (s) is the time interval in seconds.

Material Thermal Properties

Steel Properties

According to Eurocode 3 Part 1.2 [28], the thermal conductivity, specific heat, and density of structural steel have been determined. The density is taken as a constant value of 7850 kg/m³.

Figure 1 illustrates the relationship between the thermal conductivity and specific of heat of steel and temperature based on EN 1993-1-2.
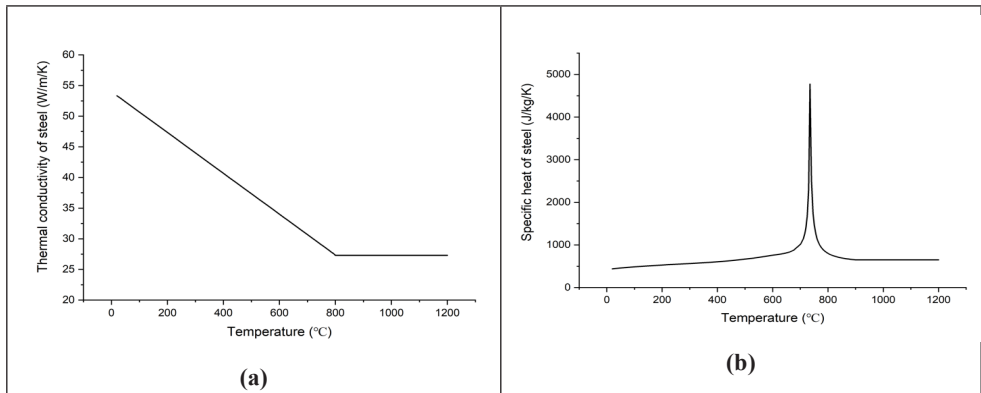
Figure 1. (a) The correlation between thermal conductivity and steel temperature (b) The variation in the specific heat of steel with temperature

Concrete Properties

Figure 2 (a) illustrates the thermal conductivity of concrete based on the concrete class (normal or light weight concrete) specified by CEN 2005c [29] while, the specific heat change of the concrete depending on the temperature is shown [29] in Figure 2 (b).
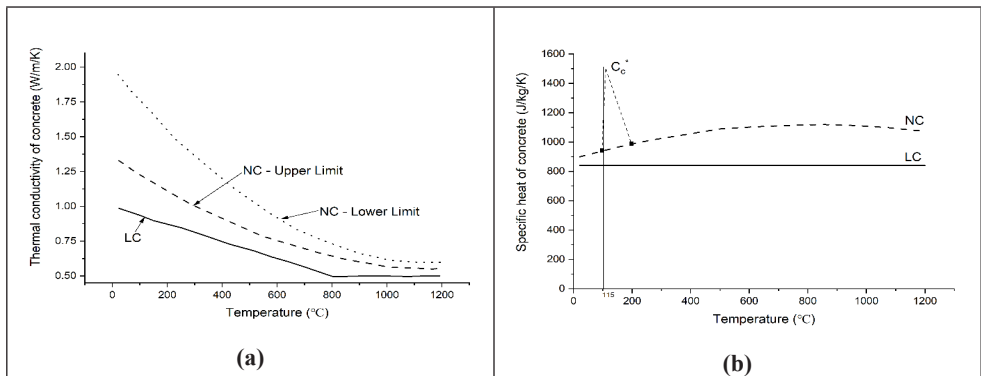

Figure 2. (a) Thermal conductivity-temperature relation of concrete (b) The change of specific heat of the concrete due to the temperature

Intumescent Coating Properties

The constant values specified in Annex E of EN 13381-8:2013 [30] were utilized. These respective values are $1000 \, J/kg \cdot K$ for the specific heat and $100 \, kg/m^3$ for the density. However, Wang, et al. [31] determined the effective thermal conductivity based on their fire tests and the ISO 834 furnace temperature. Figure 3 illustrates this effective thermal conductivity-temperature correlation. In their model, protected surfaces have been subjected to fire, which is applicable to the author's study as well.
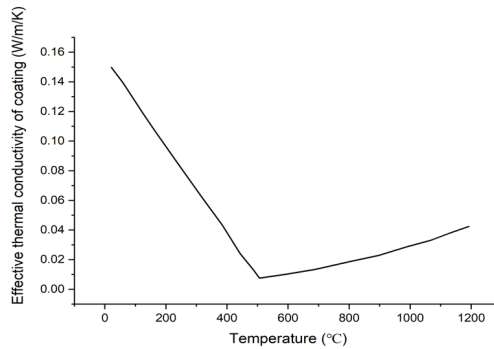
Figure 3. The relationship between effective thermal conductivity and temperature [31]

Theoretical Model Properties

In all analyses, 5.2 kN/m² as dead load (3.5 kN/m² + 1 kN/m² partitions + 0.7 kN/m² headlining, service spaces and flooring material) and 4 kN/m² as live load It is based on and $1,35G_k + 1,5Q_k$ according to Eurocode is applied for the load combination.

The properties of ATAPANEL Composite Flooring Panel (see. Figure 4) are used in the model for the sheet metal to be used in composite flooring.
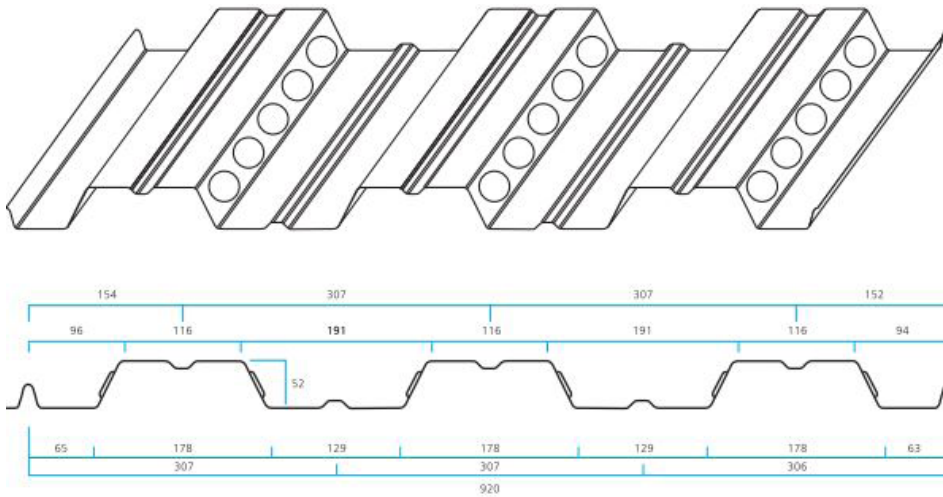
Figure 4. Features of ATAPANEL Composite Flooring Sheet

The concrete material used in the flooring is light-weight concrete and its compressive strength is C30/37 $f_{ck} = 30\ N/mm^2$).

As a result of the structural analysis, the primary beams of the steel beams used under the flooring system are IPE 360 and the secondary beams are HE 340 B, and temperature calculations for the above-mentioned variable parameters will be made

over these beams. The yield strength used in the model is $f_y = 255 \ N/mm^2$ for both beam profiles and reinforcement.

Fire Standard (ISO 834) applied in all developed mathematical models is the fire curve and 90 min. applied throughout. The results obtained (temperature calculations, etc.) are the results up to this time.

The rate of fire load applied in all developed models was taken as 90% (PD 6688-1-2:2007, Table A.2.) [27].

Microsoft Excel - Macros were used to make structural calculations and make the developed mathematical model applicable.

Intumescent paint, which is one of the passive protective systems and which swells (inflates) when exposed to heat, is used as a fire protection material. The main reasons for choosing this material are; It is aesthetic, can be applied quickly, and most importantly, it can be easily applied to complex surfaces (for example, column-beam joint areas or the area between sheet metal and steel beam in composite flooring systems).

RESULTS

Temperature predictions were performed in accordance with Section Material and Method's specifications. As a fire protection material, intumescent paint, which is one of the passive protective systems and expands when exposed to heat, is utilized. This material was chosen because it is aesthetically pleasing, can be applied fast, and, most significantly, is simple to apply to complex surfaces (for example, column-beam joint areas or the area between steel decking and steel beams in composite flooring systems).

Long and Short Span Slabs

In long-span slabs based on the mathematical model, the lengths of the primary and secondary beams are each 7.5 meters. In short-span flooring systems, the primary beam is 7.50 m long and the secondary beam is 3.75 m long. The reinforcement mesh was the same for both systems, and A252 mesh was used in accordance with BS 4483 [32]. Table 1 displays the mesh's technical characteristics.

Table 1. Properties of mesh used in the composite floor model

	Mesh sizes nominal wire spacing		Wire sizes		Cross-sectional area per meter width (m²)		Weight (1/m²)
	Main (mm)	Cross (mm)	Main (mm)	Cross (mm)	Main (mm)	Cross (mm)	kg
A252 (Ø8)	200	200	8	8	252	252	3,95

Long-Span

According to the loads indicated and the scenarios described, the neutral axis of the primary beam is in the concrete slab ($x_c = 58,2\ mm > h_1 = 52\ mm$), whereas is in the upper head section of the secondary beam ($x_c = 141,3\ mm > h_1 = 52\ mm$). Calculated bending capacities (resistances) at room temperature;

$M_{pl,Rd} = 502,2\ kNm$ (Primary beam)

$M_{pl,Rd} = 800,6\ kNm$ (Secondary beam)

The permissible deflection estimated is 160.9 mm.

Figure 5 depicts the protected and unprotected steel temperatures for the primary beam and secondary beam in the long span slab system. Protection is not applied to only one secondary beam. The results were achieved by adding 2 mm of paint to the protected profiles.
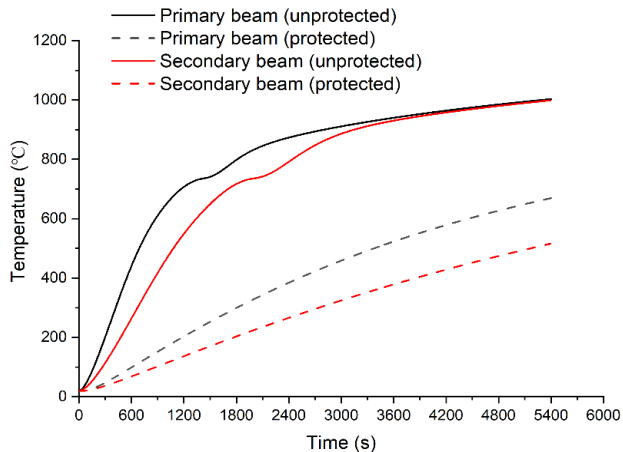
Figure 5. Protected and unprotected temperatures of primary and secondary beams

At the end of 90 minutes, the temperature of unprotected steel was 998,8 °C, and the strength loss coefficient at high temperatures is $k_{y,\theta} = 0,040$ according to Table 3.2 in Eurocode 4 Part 1.2 [29]. The temperature of the protected steel for secondary and primary beams was 515.7 °C and 669.02 °C, respectively. Figure 6 illustrates the reduction in yield strength and modulus of elasticity with temperature for steel.
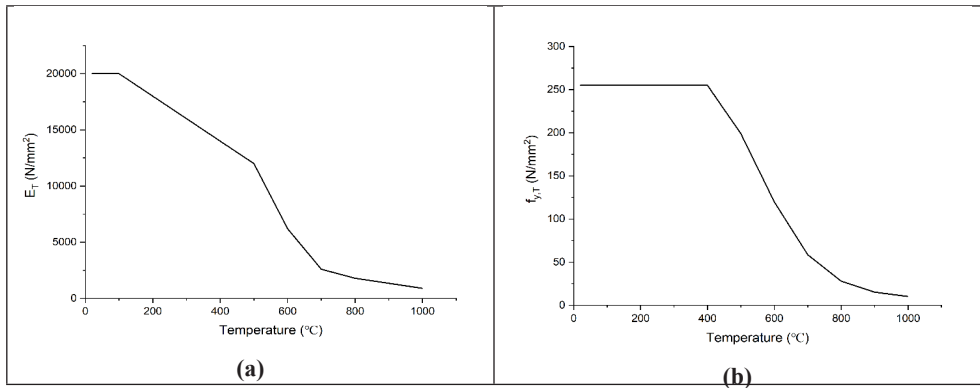
Figure 6. (a) *The change of Elasticity Modulus with temperature* (E_T) **(b)** *Yield strength-temperature variation* $(f_{y,T})$

Temperatures of reinforcement (mesh) were also determined within the scope of this study. Figure 7 displays calculated temperatures at complete, average, weighted, and continuous depths. The temperatures attained after 90 minutes are likewise displayed in Table 2. For short-term analyses, the average depth temperature is 397.1 °C, while the value is 397.1 °C for long-term analyses. The strength loss coefficient at this temperature, according to Eurocode 4 Part 1.2 Table 3.4, is 0.942.
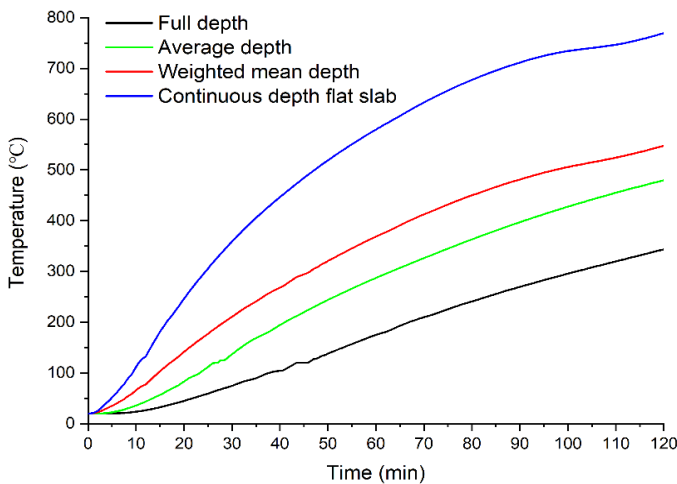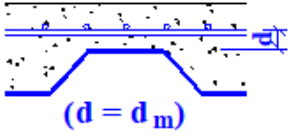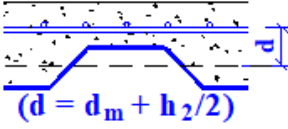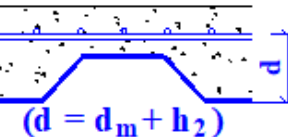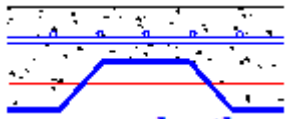
Figure 7. Temperatures of reinforcements (meshes) at various depths of concrete

Table 2. Temperatures of reinforcements (mesh) at different heights after 90 minutes in A252 (Ø8) reinforced composite flooring

Span	Temperature (°C)	Depth Locations
Continuous depth	711.6	$(d = d_m)$
Average depth	397.1	$(d = d_m + h_2/2)$
Full depth	269.6	$(d = d_m + h_2)$
Weighted mean depth	481.3	

In the case of a fire, reinforced concrete floor slabs, which are a component of the composite flooring system, will heat up from their bottom surfaces and change direction due to the temperature gradient along the thickness direction and the reduction in strength at high temperatures. Due to floor slab deformation, membrane tension increases, resulting in an increase in load bearing capacity. This circumstance was observed in this study.

Figure 8 (a) depicts the vertical displacement of the modeled composite slab. The 90-minute design produced a displacement of 181.4 mm, which exceeded the permissible displacement amount (160.9 mm). According to the study, the computation of TSlab deflection was also performed by Newman, et al. [33]. In this calculation, the slab's minimum and maximum surface temperatures as well as its dimensional parameters are considered (see Equation (3)).

The membrane action of the modeled slab results in a collapse load of 2,80 kN/m² and a load carrying capacity of 4.80 kN/m2 $(W_{p,T})$. Consequently, membrane tension contributed to an average load-bearing capacity of 1.7. The enhancement-displacement vs height curve is depicted in Figure 8 (b).

$$v = \frac{\alpha(T_2 - T_1)l^2}{19.2h} + \sqrt{\frac{0.5f_y}{E}_{t=20°C} \times \frac{3}{8}L^2}$$

(3)

v – Permissible vertical displacement

α – Coefficient of thermal expansion of concrete

T_2 - Lowest surface temperature of the flooring

T_1 - Top surface temperature of the flooring

L – The long side of the slab (long span)

l - Short side of the slab (short span)
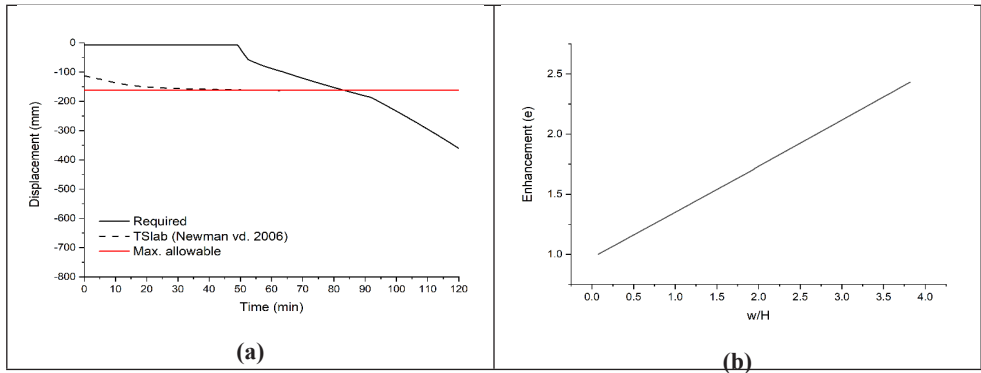

(a)   **(b)**

Figure 8. (a) Vertical displacement of slab with long-span over time (b) Enhancement (e) – Displacement/Height (w/H) result in the long-span system

Short-Span

This section details the general distinctions between the short span and the long span. Temperature calculations were determined to be unchanged as the section's dimensional parameters remained unchanged. However, as the span lengths vary, so do the bending moments and capacities.

The neutral axis is in the upper head section $(x_c = 128,7 \; mm > h_1 = 52 \; mm)$. Calculated bending capacities (resistances) at room temperature;

$M_{pl,Rd} = 424,4 \; kNm$ (Primary beam)

$M_{pl,Rd} = 800,6 \; kNm$ (Secondary beam)

The permissible deflection estimated is 104.3 mm.

Figure 9 depicts the vertical displacement of the modeled composite slab and the enhancement-displacement vs height curve for the short-span. The 90-minute design produced a displacement of 181.4 mm, which exceeded the permissible displacement amount (160.9 mm). According to the study, the computation of TSlab deflection was also performed by Newman, et al. [33]. In this calculation, the minimum and maximum surface temperatures of the slab as well as its dimensions (Eq. (3)) are taken into account.

After membrane action, the collapse load of the modeled short span slab is 4.94 kN/m², while the load carrying capacity $(W_{p,T})$ is 2.3 kN/m². Therefore, membrane

tension had no effect on the load carrying capability. Therefore, no membrane development was detected in the short span beam. The enhancement-displacement vs height curve for the short-span is presented in Figure 9.
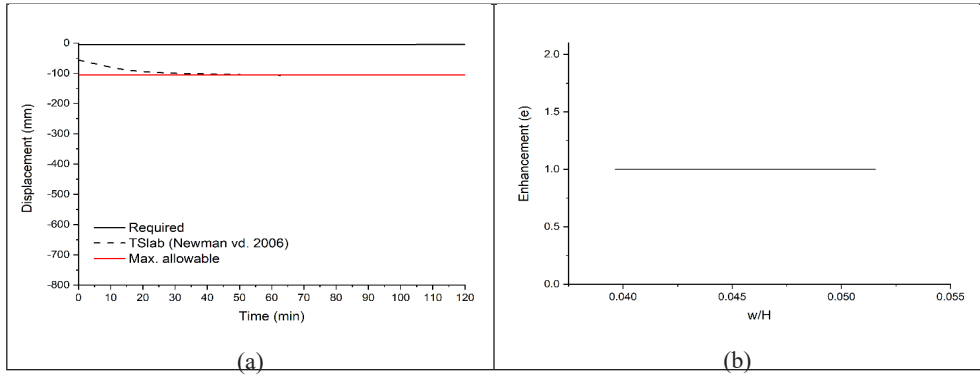

(a)   (b)

Figure 9. (a) Vertical displacement of slab with short-span over time (b) Enhancement "(e)" – Displacement/Height (w/H) result in the short-span system

All Sections (Primary and Secondary Beams) Fully Protected

Long span slabs were chosen and fire protection (intumescent paint) is applied with 2 mm thickness to all primary and secondary beams in the model. This section discusses the key distinctions.

Since the cross-sections of the primary beam (IPE 360) and secondary beam (HE 340 B) are identical, the protected and unprotected steel temperature calculations and the change of mechanical properties $(E_T, f_{y,T})$ with temperature are identical to those reported in the previous section. Figure 10 depicts the vertical displacement of the modeled composite slab with all protected beams and the enhancement-displacement vs height curve for the short-span. Nevertheless, the moment capacity of the composite flooring has increased as a result of the application of fire protection to all beams. The collapse load of the slab modeled in this section is 2.80 kN/m², whereas its load-bearing capacity $(W_{p,T})$ is 6 kN/m². In this instance, it contributed to an average coefficient of load carrying capacity of 2.1.

Although the allowable amount of deflection was 160,9 mm, there was no collapse due to the action of fire protection despite the deflection of 290.4 mm after 90 minutes. This demonstrated that the passive protection (paint layer) substantially improved the composite flooring's overall fire resistance.
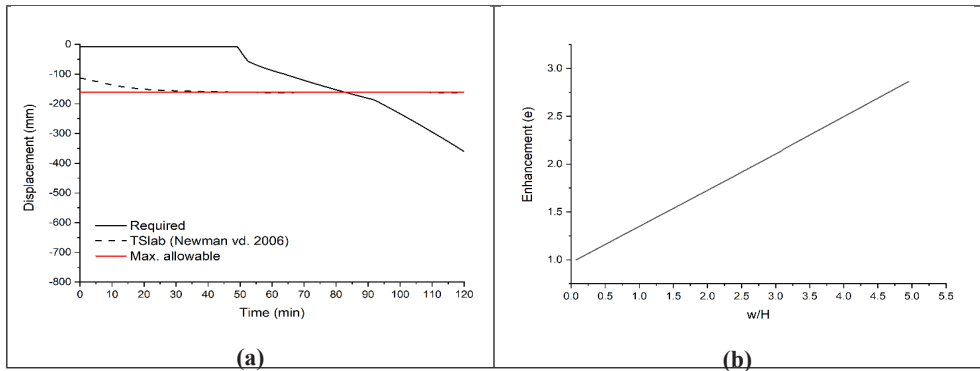
Figure 10. (a) Vertical displacement of entirely fire-protected slab with long-span (b) Enhancement "(e)" – Displacement/Height (w/H) result

CONCLUSION

In this paper, various mathematical models for a composite slab with known beam dimensions have been built and analyzed in actual dimensions. The fire curve in these produced mathematical models is the ISO 834 standard fire curve.

The following is a concise summary of the conclusions made within the scope of this research study:

- Since the amount of displacement (vertical displacement) in long-span slabs will be significantly more than in short span slabs, membrane tensile action has occurred. This circumstance improves the load-carrying capacity considerably. The development of a membrane state was not observed in short-span slabs.

- The application of a 2 mm fireproof paint coating contributed to the fire performance of the steel beams (main and secondary beams with one secondary beam unprotected), and the temperatures obtained at the end of the 90-minute design period were significantly lower than the temperatures of the unprotected steel.

- Predictions of strength and modulus of elasticity at temperatures of unprotected steel profiles were made. At the end of the fire interval, both their original strength (S255) and elasticity (200 GPA) modules had dropped to 4%.

- It was determined that applying full protection to all beams (main and secondary) significantly enhances the fire performance of the composite flooring in comparison to the unprotected state of one secondary beam. In addition, membrane development increased in these slabs compared to the unprotected secondary beam. In addition, the load-carrying capacity has significantly increased.

REFERENCES

[1]   A. S. o. C. E. ASCE, Standard for the Structural Design of Composite Slabs and Standard Practice for Construction and Inspection of Composite Slabs (Standard for the Structural Design of Composite Slabs and Standard Practice for Construction and Inspection of Composite Slabs). 1994.

[2]   B. S. I. (BSI), "British Standard BS 5950, Structural Use of Steelwork in Buildings, Part 8: Code of Practice for Fire Resistant Design," ed. London, 1990.

[3]   CEN, "EN 1993-1-1: Eurocode 3. Design of Steel Structures," in *Part 1-1: General rules and rules for buildings*, ed. BSI: London, 2005.

[4]   C. Yorgun, Çelik Sac-Beton Kompozit Döşeme Sistemlerinin Uygulamalarına Yönelik Değerlendirmeler, TÜRKIYE MÜHENDISLIK HABERLERI. 2005, 435: 60-64

[5]   J. Mahachi, "A comparison of two decking profiles subjected to fatigue load," in *RILEM International Conference on Dynamic Behaviour of Concrete Structures*, Bratislava, Slovakia, 1995, pp. 210-211.

[6]   J. Mahachi, "Response of composite bond-deck slabs to fatigue load.," in *5th International Conference on Steel Structures*, Jakarta, Indonesia, 1994, pp. 177-182.

[7]   C. Bailey and D. B. Moore, The structural behaviour of steel frames with composite floorslabs subject to fire: Part 1: Theory, The Structural engineer. 2000, 78: 19-27

[8]   C. Üstündağ and O. C. Çelik, Çok Katlı Çelik Yapılarda Kullanılan Kompozit Döşeme Sistemleri, Yapı Dergisi. 2006, 293: 88-92

[9]   S. Yılmaz, "Kompozit Döşeme Elemanlarda Pozitif Moment Taşıma Kapasitesinin İrdelenmesi," MSc, İstanbul Teknik Üniversitesi, Fen Bilimleri Enstitüsü, 2008.

[10]   L. C. S. Cordeiro and V. P. Silva, Sobre o dimensionamento de laje mista de aço e concreto em situação de incêndio, Revista da Estrutura de Aço. 2016, 5(1): 39-58

[11]   S. Selamet and T. F. Yolaçan, Çelik-Beton Kompozit Kat Döşemesi Yangın Dayanım Deneyi, Teknik Dergi. 2017: 8007-8022

[12]   G.-Q. Li and W.-Y. Wang, A simplified approach for fire-resistance design of steel-concrete composite beams, Steel and Composite Structures. 2013, 14(3): 295-312

[13]   T. T. Nguyen, K. H. Tan, and I. W. Burgess, Behaviour of composite slab-beam systems at elevated temperatures: Experimental and numerical investigation, Engineering Structures. 2015, 82: 199-213

[14]   B. K. Cirpici, S. N. Orhan, and T. Kotan, "Numerical Modelling of Heat Transfer Through Protected Composite Structural Members," presented at the International Civil Engineering and Architecture Conference 2019 (ICEARC 2019), Trabzon-Turkey, 2019.

[15]   B. K. Cirpici, S. N. Orhan, and T. Kotan, Numerical modelling of heat transfer through protected composite structural members, Challenge Journal of Structural Mechanics. 2019, 5(3): 96-107

[16]   B. K. Cirpici, S. N. Orhan, and T. Kotan, "Thermal performance of protected composite slab-beam systems exposed to fire," in *3rd International Conference on Advanced Engineering Technologies*, Bayburt-Turkey, 2019.

[17]   B. K. Cirpici, S. N. Orhan, and T. Kotan, Finite element study on composite slab-beam systems under various fire exposures, Steel and Composite Structures. 2020, 37(5): 589-603

[18]   Y. Wang *et al.*, Post-fire behaviour of continuous reinforced concrete slabs under different fire conditions, Engineering Structures. 2021, 226

[19]   T. T. Nguyen and K. H. Tan, Behaviour of composite floors with different sizes of edge beams in fire, Journal of Constructional Steel Research. 2017, 129: 28-41

[20]   D. Pantousa and E. Mistakidis, Rotational capacity of pre-damaged I-section steel beams at elevated temperatures, Steel and Composite Structures. 2017, 23(1): 53-66

[21]   N. Alam, A. Nadjai, F. Ali, and W. Nadjai, Structural response of unprotected and protected slim floors in fire, Journal of Constructional Steel Research. 2018, 142: 44-54

[22] J. Jiang, J. A. Main, J. M. Weigand, and F. H. Sadek, Thermal performance of composite slabs with profiled steel decking exposed to fire effects, Fire Safety Journal. 2018, 95: 25-41

[23] L. Lim and C. Wade, "Experimental Fire Tests of Two-Way Concrete Slabs," in "Fire Engineering Research Report 02/12," University of Canterbury, Porirua City, New Zealand, 2002.

[24] P. A. G. Piloto, C. Balsa, F. Ribeiro, and R. Rigobello, Computational Simulation of the Thermal Effects on Composite Slabs Under Fire Conditions, Mathematics in Computer Science. 2020,

[25] P. A. G. Piloto, C. Balsa, F. F. Ribeiro, and R. Rigobello. *Three-dimensional numerical analysis on the fire behaviour of composite slabs with steel deck, Lecture Notes in Civil Engineering*, vol. 1, pp. 12-30, 2020.

[26] P. A. G. Piloto, C. Balsa, L. M. C. Santos, and É. F. A. Kimura, Effect of the load level on the resistance of composite slabs with steel decking under fire conditions, Journal of Fire Sciences. 2020, 38(2): 212-231

[27] CEN, "EN 1991-1-2:2002 Eurocode 1: Actions on structures," in *Part 1.2: General actions - Actions on structures exposed to fire*, ed. BSI: London, 2002.

[28] CEN, "EN 1993-1-2: Eurocode 3. Design of Steel Structures," in *Part 1.2: General Rules - Structural fire design*, ed. BSI: London, 2005.

[29] CEN, "EN 1994-1-2:2005, Eurocode 4: Design of Composite Steel and Concrete Structures – Part 1-2: General Rules – Structural Fire Design," in *Part 1-2: General Rules – Structural Fire Design*, ed. BSI: London, 2005.

[30] CEN, "EN 13381-8:2013 Test methods for determining the contribution to the fire resistance of structural members.," in *Part 8: Applied reactive protection to steel members*, ed. BSI: London, 2013.

[31] L. L. Wang, Y. C. Wang, J. F. Yuan, and G. Q. Li, Thermal conductivity of intumescent coating char after accelerated aging, Fire and Materials. 2013, 37(6): 440-456

[32] B. S. I. (BSI), "British Standard BS 4483:2005, Steel fabric for the reinforcement of concrete - Specification," ed. London, 2005.

[33] G. M. Newman, J. T. Robinson, and C. G. Bailey, "Fire safe design: A new approach to multi-storey steel-framed buildings," ed. UK: The Steel Construction Institute, 2006.